(FILE 'HOME' ENTERED AT 15:05:11 ON 30 NOV 2000)

FILE 'REGISTRY' ENTERED AT 15:05:52 ON 30 NOV 2000

```
             E RECA
L1          347 S E3
```

INDEX 'ADISALERTS, ADISINSIGHT, AGRICOLA, AIDSLINE, ANABSTR, AQUASCI, BIOBUSINESS, BIOCOMMERCE, BIOSIS, BIOTECHABS, BIOTECHDS, BIOTECHNO, CABA, CANCERLIT, CAPLUS, CEABA, CEN, CIN, CONFSCI, CROPB, CROPU, DDFB, DDFU, DGENE, DRUGB, DRUGLAUNCH, DRUGMONOG2, ...' ENTERED AT 15:18:35 ON 30 NOV 2000

```
             SEA RECA
            ---------
              3   FILE ADISALERTS
              1   FILE ADISINSIGHT
            146   FILE AGRICOLA
              9   FILE AIDSLINE
              1   FILE ANABSTR
             15   FILE AQUASCI
             39   FILE BIOBUSINESS
             15   FILE BIOCOMMERCE
           3663   FILE BIOSIS
            499   FILE BIOTECHABS
            499   FILE BIOTECHDS
           2157   FILE BIOTECHNO
            152   FILE CABA
            570   FILE CANCERLIT
           4294   FILE CAPLUS
             54   FILE CEABA
              1   FILE CEN
              7   FILE CIN
             74   FILE CONFSCI
              5   FILE CROPU
              6   FILE DDFB
             35   FILE DDFU
            399   FILE DGENE
              6   FILE DRUGB
              2   FILE DRUGLAUNCH
              4   FILE DRUGMONOG2
             72   FILE DRUGU
             17   FILE EMBAL
           2692   FILE EMBASE
            952   FILE ESBIOBASE
              3   FILE FROSTI
             54   FILE FSTA
           1124   FILE GENBANK
             10   FILE HEALSAFE
             74   FILE IFIPAT
            154   FILE JICST-EPLUS
              3   FILE KOSMET
           2087   FILE LIFESCI
           3361   FILE MEDLINE
             40   FILE NIOSHTIC
             59   FILE NTIS
              2   FILE OCEAN
              4   FILE PHIN
             67   FILE PROMT
           2969   FILE SCISEARCH
           1819   FILE TOXLINE
           2159   FILE TOXLIT
           1037   FILE USPATFULL
             78   FILE WPIDS
             78   FILE WPINDEX
L2             QUE RECA
              ---------
```

FILE 'CAPLUS, BIOSIS, MEDLINE, SCISEARCH, EMBASE, TOXLIT, BIOTECHNO,

LIFESCI, TOXLINE, GENBANK, USPATFULL, ESBIOBASE, CANCERLIT, BIOTECHDS, DGENE, JICST-EPLUS, CABA, AGRICOLA, WPIDS, CONFSCI, IFIPAT, DRUGU, PROMT, NTIS, CEABA, FSTA, NIOSHTIC, BIOBUSINESS, ...' ENTERED AT 15:18:57 ON 30 NOV 2000

L3 30953 S L2
L4 9955 S L3 AND E COLI
L5 484 S L4 AND MUTANT AND ATP
L6 395 DUP REM L5 (89 DUPLICATES REMOVED)
L7 0 S L5 AND MAW AND MOTIF
L8 0 S L5 AND MAW
L9 0 S L5 AND MAKES ATP WORK

FILE 'REGISTRY' ENTERED AT 15:39:14 ON 30 NOV 2000
L10 0 S RECA AND E COLI

L11 ANSWER 1 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 290393-97-4 REGISTRY
CN **RecA protein (Vibrio cholerae strain N16961 gene VC0543) (9CI)**
(CA INDEX NAME)
OTHER NAMES:
CN GenBank AE004140-derived protein GI 9654971
FS PROTEIN SEQUENCE
MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

=> d 1-53

L11 ANSWER 1 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 290393-97-4 REGISTRY
CN **RecA protein (Vibrio cholerae strain N16961 gene VC0543) (9CI)**
(CA INDEX NAME)
OTHER NAMES:
CN GenBank AE004140-derived protein GI 9654971
FS PROTEIN SEQUENCE
MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 2 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 263000-78-8 REGISTRY
CN **RecA protein (Neisseria meningitidis strain Z2491 gene recA) (9CI)**
(CA INDEX NAME)
OTHER NAMES:
CN GenBank AL162756-derived protein GI 7380300
FS PROTEIN SEQUENCE
MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 3 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 261933-60-2 REGISTRY
CN **RecA protein (Chlamydia muridarum strain Nigg gene TC0019) (9CI)**
(CA INDEX NAME)
OTHER NAMES:
CN GenBank AE002270-derived protein GI 7190059
FS PROTEIN SEQUENCE
MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 4 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 260036-57-5 REGISTRY
CN **RecA protein (Neisseria meningitidis strain MD58 gene NMB1445) (9CI)** (CA INDEX NAME)
FS PROTEIN SEQUENCE
MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 5 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 257275-98-2 REGISTRY
CN **RecA protein (Campylobacter jejuni strain NCTC 11168 gene recA) (9CI)** (CA INDEX NAME)
OTHER NAMES:
CN GenBank AL139079-derived protein GI 6969089
FS PROTEIN SEQUENCE
MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 6 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 253842-10-3 REGISTRY
CN **RecA protein (Arabidopsis thaliana gene At2g19490) (9CI)** (CA INDEX NAME)
FS PROTEIN SEQUENCE
MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 7 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 250319-20-1 REGISTRY
CN **RecA protein (Deinococcus radiodurans strain R1 gene DR2340) (9CI)** (CA INDEX NAME)
OTHER NAMES:
CN Enzyme (Deinococcus radiodurans strain KD8301 clone pJTM89 and pJTM50 gene recA)
CN GenBank AB005471-derived protein GI 2251089
CN GenBank AE002065-derived protein GI 6460153
CN **RecA protein (Deinococcus radiodurans strain KD8301 clone pJTM89 orf105c)**
FS PROTEIN SEQUENCE
MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS, TOXLIT

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
2 REFERENCES IN FILE CA (1967 TO DATE)
2 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 8 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 235426-73-0 REGISTRY
CN **RecA protein (Rhodopseudomonas palustris strain No.7 gene recA) (9CI)** (CA INDEX NAME)
OTHER NAMES:
CN GenBank D84467-derived protein GI 5103048
FS PROTEIN SEQUENCE
MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS, TOXLIT

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 9 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 225225-84-3 REGISTRY
CN **RecA protein (Vibrio cholerae strain M549 gene recA C-terminal fragment) (9CI)** (CA INDEX NAME)
OTHER NAMES:
CN GenBank AF117881-derived protein GI 4566895
FS PROTEIN SEQUENCE
MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS, TOXLIT

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 10 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 225225-83-2 REGISTRY
CN **RecA protein (Vibrio cholerae strain M793 gene recA C-terminal fragment) (9CI)** (CA INDEX NAME)
OTHER NAMES:
CN GenBank AF117878-derived protein GI 4566892
CN GenBank AF117879-derived protein GI 4566893
CN GenBank AF117880-derived protein GI 4566894
CN **RecA protein (Vibrio cholerae strain M794 gene recA C-terminal fragment)**
CN **RecA protein (Vibrio cholerae strain M967 gene recA C-terminal fragment)**
FS PROTEIN SEQUENCE
MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS, TOXLIT

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 11 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 223708-10-9 REGISTRY
CN RecA recombination protein (Chlamydia pneumoniae gene recA) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN GenBank AE001658-derived protein GI 4377070

CN GenBank AE002267-derived protein GI 7190022
CN **RecA protein (Chlamydophila pneumoniae AR39 strain AR39 gene CP1110)**
FS PROTEIN SEQUENCE
MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
2 REFERENCES IN FILE CA (1967 TO DATE)
2 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 12 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 220916-57-4 REGISTRY
CN DNA (Thermus aquaticus gene recA fragment) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN **DNA (Thermus aquaticus gene recA RecA protein fragment)**
FS NUCLEIC ACID SEQUENCE
MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES TO NON-SPECIFIC DERIVATIVES IN FILE CA
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 13 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 220797-11-5 REGISTRY
CN **RecA protein (Streptococcus parasanguis clone pVT1356 gene recA) (9CI)** (CA INDEX NAME)
OTHER NAMES:
CN GenBank AF069745-derived protein GI 3212104
CN **RecA protein (Streptococcus parasanguis strain FW213 clone pVT1356 gene recA)**
FS PROTEIN SEQUENCE
MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS, TOXLIT

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 14 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 216497-67-5 REGISTRY
CN **Protein (Yersinia pestis virulence plasmid pMT1 RecA protein-like) (9CI)** (CA INDEX NAME)
OTHER NAMES:
CN GenBank AF053947-derived protein GI 2996328
CN GenBank AF074611-derived protein GI 3883111
CN Recombinase (plasmid pMT1 open reading frame 26 RecA-like)
FS PROTEIN SEQUENCE
MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS, TOXLIT

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
2 REFERENCES IN FILE CA (1967 TO DATE)
2 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 15 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 215107-47-4 REGISTRY
CN **RecA protein (Chlamydia trachomatis gene recA) (9CI)** (CA INDEX NAME)
OTHER NAMES:
CN GenBank AE001335-derived protein GI 3329099
FS PROTEIN SEQUENCE
MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 16 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 211430-71-6 REGISTRY
CN **642-693-RecA protein [693-alanine] (Mycobacterium tuberculosis gene recA) (9CI)** (CA INDEX NAME)
FS PROTEIN SEQUENCE
MF C267 H429 N75 O81 S2
CI MAN
SR CA
LC STN Files: CA, CAPLUS

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 17 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 210411-08-8 REGISTRY
CN **RecA protein (Blastochloris viridis strain DSM133) (9CI)** (CA INDEX NAME)
FS PROTEIN SEQUENCE
MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS, TOXLIT

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 18 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 209675-57-0 REGISTRY
CN **RecA protein (recA) (Treponema pallidum gene TP0692) (9CI)** (CA INDEX NAME)
OTHER NAMES:
CN GenBank AE001243-derived protein GI 3322995
FS PROTEIN SEQUENCE
MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 19 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 208734-23-0 REGISTRY
CN **RecA protein (Chlorobium tepidum) (9CI)** (CA INDEX NAME)
OTHER NAMES:
CN GenBank AF037258-derived protein GI 2760912

FS PROTEIN SEQUENCE
MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS, TOXLIT

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 20 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 208734-22-9 REGISTRY
CN **RecA protein (Chloroflexus aurantiacus strain J-10-f1) (9CI)**
(CA INDEX NAME)
OTHER NAMES:
CN GenBank AF037259-derived protein GI 2760914
FS PROTEIN SEQUENCE
MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS, TOXLIT

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 21 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 208129-44-6 REGISTRY
CN **RecA protein (human gene HsR51h3 recombination-repair) (9CI)**
(CA INDEX NAME)
OTHER NAMES:
CN GenBank Y15572-derived protein GI 3005965
FS PROTEIN SEQUENCE
MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 22 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 208129-43-5 REGISTRY
CN **RecA protein (human gene HsR51h2 recombination-repair) (9CI)**
(CA INDEX NAME)
OTHER NAMES:
CN GenBank Y15571-derived protein GI 3005963
FS PROTEIN SEQUENCE
MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 23 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 208129-42-4 REGISTRY
CN **RecA protein (mouse gene HsR51h3 recombination-repair) (9CI)**
(CA INDEX NAME)
OTHER NAMES:
CN GenBank AB007040-derived protein GI 2696715
CN GenBank AF034955-derived protein GI 2920580

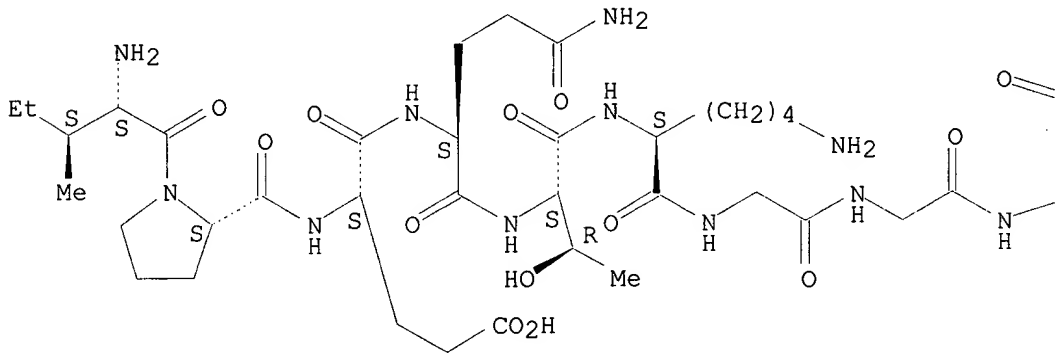
CN GenBank Y15570-derived protein GI 3005974
CN Protein (mouse gene Rad51d)
CN Trad (Mus musculus IMAGE clone 746795 gene Trad)
FS PROTEIN SEQUENCE
MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS, TOXLIT

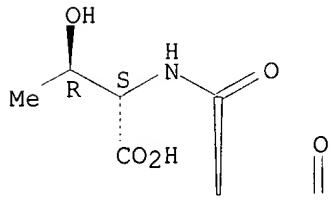
*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
3 REFERENCES IN FILE CA (1967 TO DATE)
3 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 24 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 207919-94-6 REGISTRY
CN L-Threonine, L-isoleucyl-L-prolyl-L-.alpha.-glutamyl-L-glutaminyl-L-threonyl-L-lysylglycylglycyl-L-arginyl-L-asparaginyl-L-threonyl-L-methionyl-L-asparaginyl-L-valyl-L-phenylalanylglycyl-L-methionylglycyl-L-isoleucyl- (9CI) (CA INDEX NAME)
OTHER NAMES:
CN **193-212-RecA protein scrambled peptide WT-Scr (Escherichia coli DNA-binding loop L2)**
FS PROTEIN SEQUENCE; STEREOSEARCH
MF C91 H151 N27 O29 S2
SR CA
LC STN Files: CA, CAPLUS

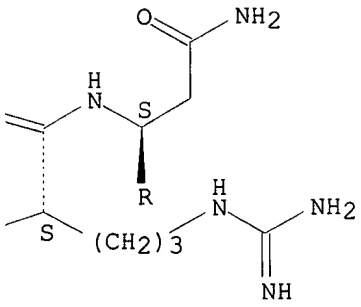
Absolute stereochemistry.

PAGE 1-A

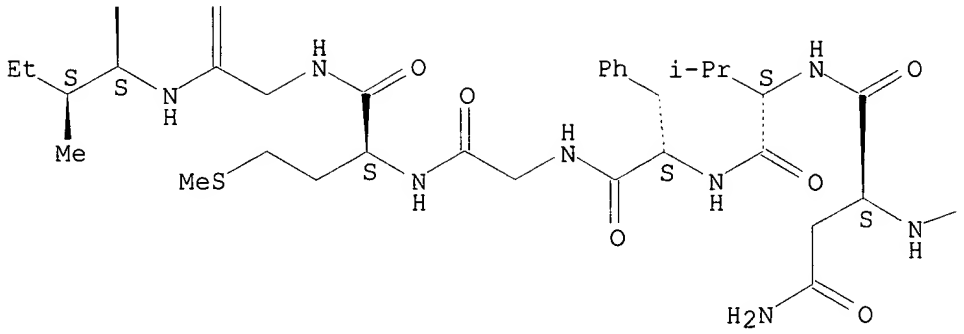


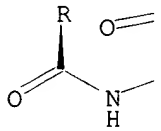


PAGE 1-B

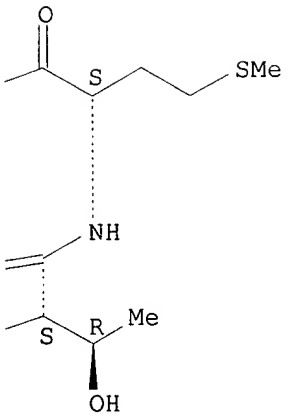


PAGE 2-A





PAGE 2-B



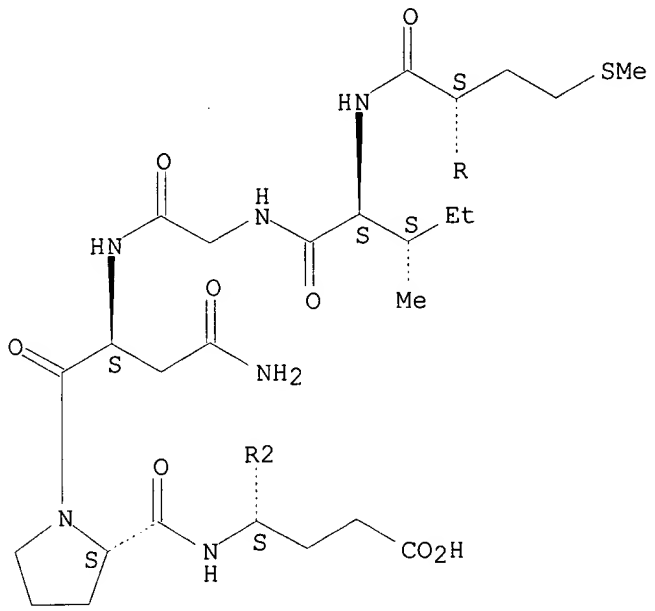
1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 25 OF 53 REGISTRY COPYRIGHT 2000 ACS

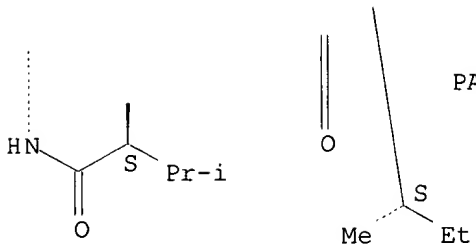
RN 207919-93-5 REGISTRY
CN Glycine, L-asparaginyl-L-glutaminyl-L-isoleucyl-L-arginyl-L-methionyl-L-lysyl-L-isoleucylglycyl-L-valyl-L-methionyl-L-isoleucylglycyl-L-asparaginyl-L-prolyl-L-.alpha.-glutamyl-L-threonyl-L-threonyl-L-threonylglycyl- (9CI) (CA INDEX NAME)
OTHER NAMES:
CN **193-212-RecA protein peptide IECO (Escherichia coli DNA-binding loop L2)**
FS PROTEIN SEQUENCE; STEREOSEARCH
MF C88 H153 N27 O29 S2
SR CA
LC STN Files: CA, CAPLUS

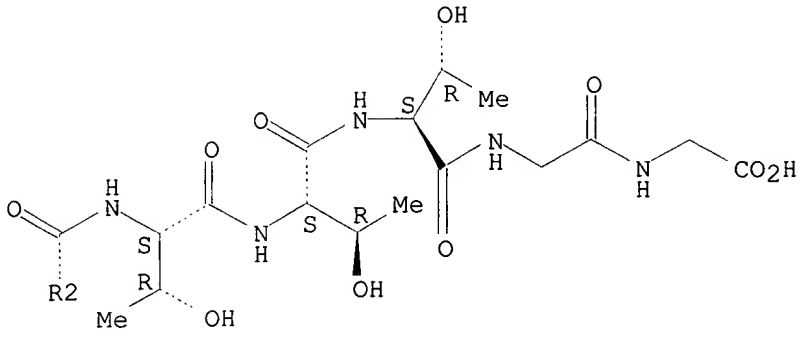
Absolute stereochemistry.

PAGE 1-A



* STRUCTURE DIAGRAM TOO LARGE FOR DISPLAY - AVAILABLE VIA OFFLINE PRINT *



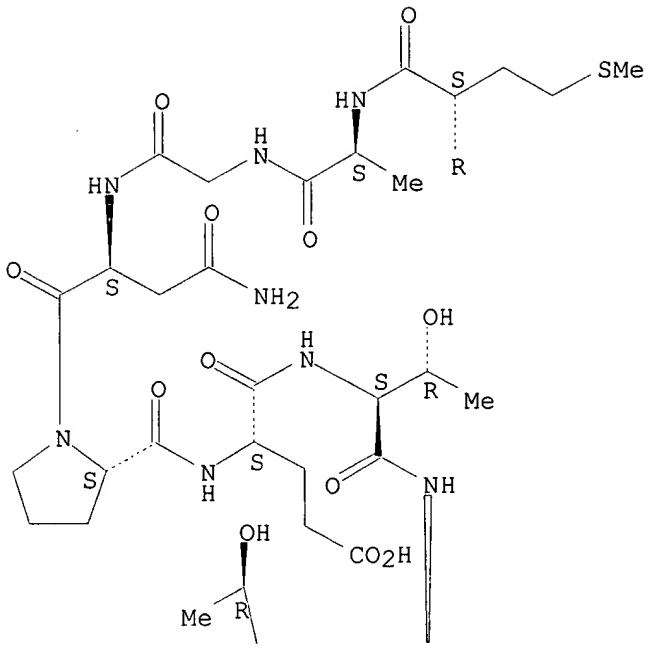


1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 26 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 207919-92-4 REGISTRY
CN Glycine, L-asparaginyl-L-glutaminyl-L-isoleucyl-L-arginyl-L-methionyl-L-lysyl-L-isoleucylglycyl-L-valyl-L-methionyl-L-alanylglycyl-L-asparaginyl-L-prolyl-L-.alpha.-glutamyl-L-threonyl-L-threonyl-L-threonylglycyl- (9CI) (CA INDEX NAME)
OTHER NAMES:
CN **193-212-RecA protein peptide AECO (Escherichia coli DNA-binding loop L2)**
FS PROTEIN SEQUENCE; STEREOSEARCH
MF C85 H147 N27 O29 S2
SR CA
LC STN Files: CA, CAPLUS

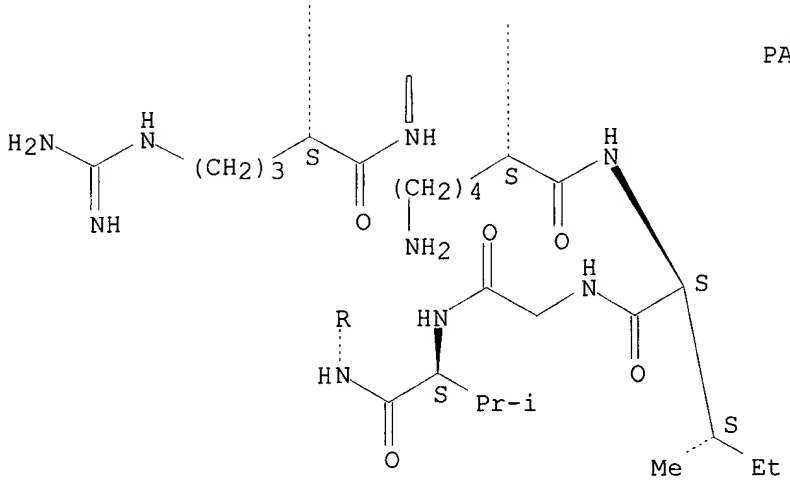
Absolute stereochemistry.

PAGE 1-A



* STRUCTURE DIAGRAM TOO LARGE FOR DISPLAY - AVAILABLE VIA OFFLINE PRINT *

PAGE 3-A



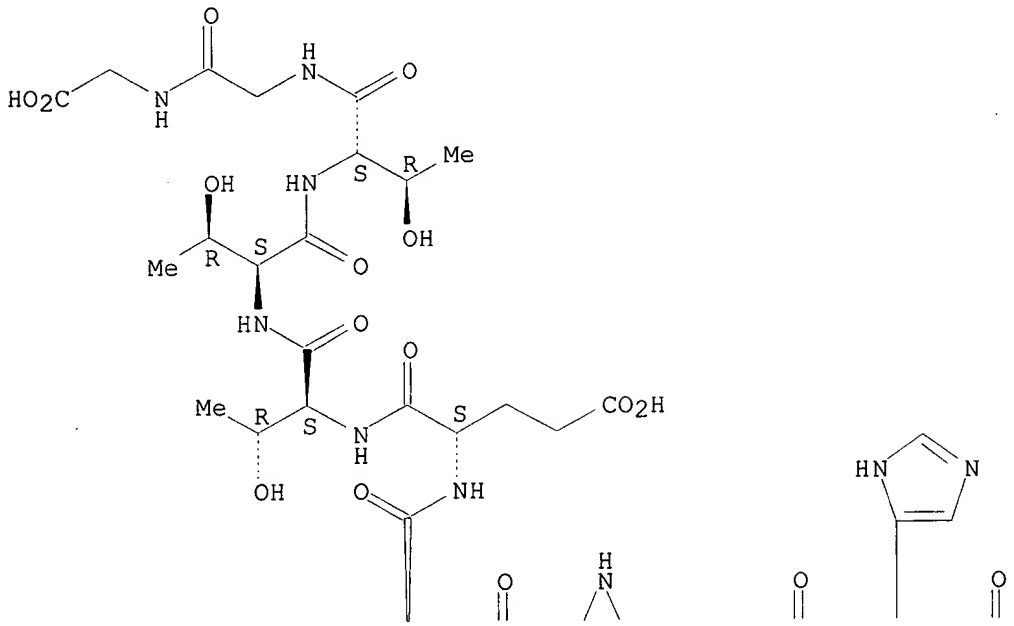
1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 27 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 207919-91-3 REGISTRY
CN Glycine, L-asparaginyl-L-glutaminyl-L-isoleucyl-L-arginyl-L-methionyl-L-lysyl-L-isoleucylglycyl-L-valyl-L-methionyl-L-histidylglycyl-L-asparaginyl-L-prolyl-L-.alpha.-glutamyl-L-threonyl-L-threonyl-L-threonylglycyl- (9CI) (CA INDEX NAME)
OTHER NAMES:
CN **193-212-RecA protein peptide HECO (Escherichia coli DNA-binding loop L2)**
FS PROTEIN SEQUENCE; STEREOSEARCH
MF C88 H149 N29 O29 S2

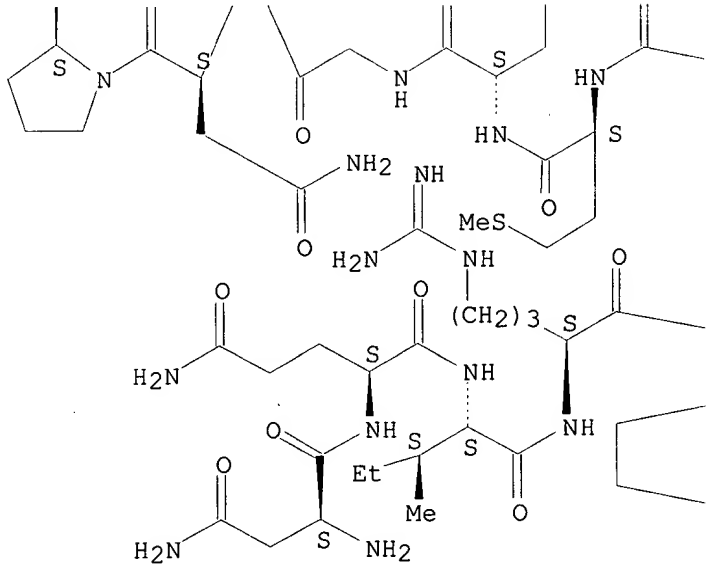
SR CA
LC STN Files: CA, CAPLUS

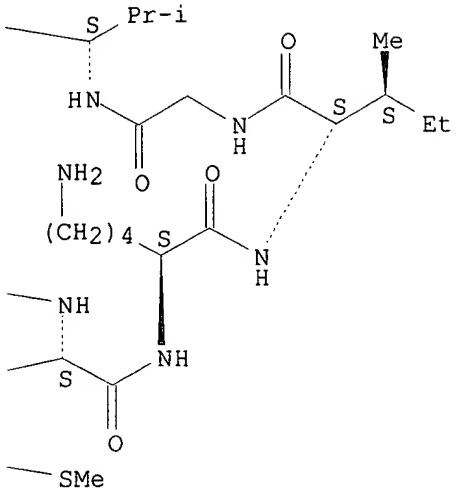
Absolute stereochemistry.

PAGE 1-A



PAGE 2-A





1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 28 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 207870-91-5 REGISTRY
CN **RecA protein (Gluconobacter oxydans clone pGO1 gene recA) (9CI)** (CA INDEX NAME)
OTHER NAMES:
CN GenBank U21001-derived protein GI 915212
FS PROTEIN SEQUENCE
MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS, TOXLIT

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 29 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 206234-02-8 REGISTRY
CN **DNA (mouse gene HsR51h3 recombination-repair RecA protein cDNA plus flanks) (9CI)** (CA INDEX NAME)
OTHER NAMES:
CN GenBank Y15570
FS NUCLEIC ACID SEQUENCE
MF Unspecified
CI MAN
SR GenBank
LC STN Files: CA, CAPLUS, GENBANK

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 30 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 206234-00-6 REGISTRY
CN **DNA (human gene HsR51h3 recombination-repair RecA protein cDNA plus flanks) (9CI)** (CA INDEX NAME)
OTHER NAMES:
CN GenBank Y15572
FS NUCLEIC ACID SEQUENCE
MF Unspecified
CI MAN

SR GenBank
LC STN Files: CA, CAPLUS, GENBANK

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 31 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 206233-98-9 REGISTRY
CN **DNA (human gene HsR51h2 recombination-repair RecA protein cDNA plus flanks) (9CI)** (CA INDEX NAME)
OTHER NAMES:
CN GenBank Y15571
FS NUCLEIC ACID SEQUENCE
MF Unspecified
CI MAN
SR GenBank
LC STN Files: CA, CAPLUS, GENBANK

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 32 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 204935-34-2 REGISTRY
CN **RecA protein (Burkholderia cepacia clone pPM3.6 gene recA) (9CI)** (CA INDEX NAME)
OTHER NAMES:
CN GenBank U70431-derived protein GI 1575711
FS PROTEIN SEQUENCE
MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS, TOXLIT

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 33 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 203673-88-5 REGISTRY
CN **RecA protein (Xanthomonas oryzae oryzae clone pSM-A1 gene recA) (9CI)** (CA INDEX NAME)
OTHER NAMES:
CN GenBank AF013600-derived protein GI 2460018
FS PROTEIN SEQUENCE
MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS, TOXLIT

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 34 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 200821-01-8 REGISTRY
CN **RecA protein (recA) (Borrelia burgdorferi strain B31 gene BB0131) (9CI)** (CA INDEX NAME)
OTHER NAMES:
CN GenBank AE001124-derived protein GI 2688005
FS PROTEIN SEQUENCE
MF Unspecified
CI MAN
SR CA

LC STN Files: CA, CAPLUS

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

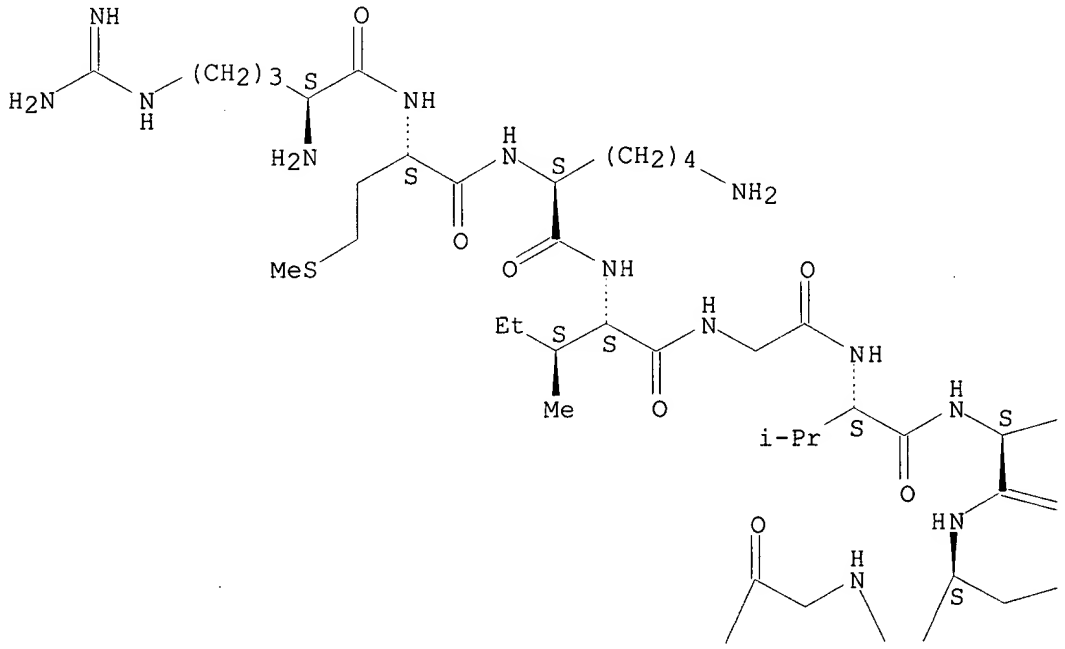
L11 ANSWER 35 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 198908-75-7 REGISTRY
CN **RecA protein (Mycobacterium tuberculosis strain So93 gene recA) (9CI)** (CA INDEX NAME)
OTHER NAMES:
CN GenBank AJ000011-derived protein GI 2598002
FS PROTEIN SEQUENCE
MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS

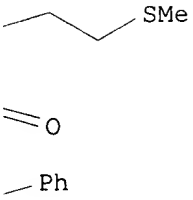
*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 36 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 198828-10-3 REGISTRY
CN L-Threonine, L-arginyl-L-methionyl-L-lysyl-L-isoleucylglycyl-L-valyl-L-methionyl-L-phenylalanylglycyl-L-asparaginyl-L-prolyl-L-.alpha.-glutamyl-L-threonyl- (9CI) (CA INDEX NAME)
OTHER NAMES:
CN **196-209-RecA protein peptide WT-14 (Escherichia coli DNA-binding loop L2)**
FS PROTEIN SEQUENCE; STEREOSEARCH
MF C68 H113 N19 O20 S2
SR CAS Registry Services
LC STN Files: CA, CAPLUS

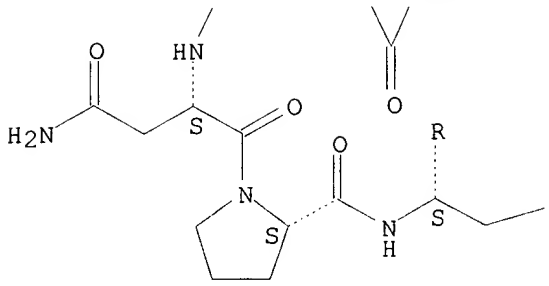
Absolute stereochemistry.

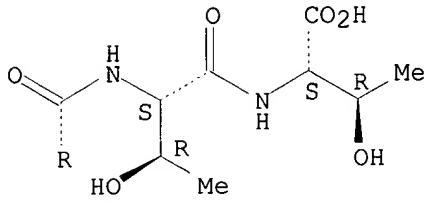
PAGE 1-A



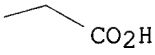


PAGE 2-A





PAGE 2-B



1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 37 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 198433-34-0 REGISTRY
CN **RecA protein (Bacteroides thetaiotaomicron strain 5482 gene recA fragment) (9CI)** (CA INDEX NAME)
OTHER NAMES:
CN GenBank U63514-derived protein GI 1480909
FS PROTEIN SEQUENCE
MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS, TOXLIT

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 38 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 190338-99-9 REGISTRY
CN **RecA protein (Spirulina platensis strain IAM-M135 gene recA reduced) (9CI)** (CA INDEX NAME)
OTHER NAMES:
CN GenBank U33924-derived protein GI 976445
CN Recombination protein A (Spirulina platensis strain IAM-M135 gene recA reduced)
FS PROTEIN SEQUENCE
MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS, TOXLIT

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 39 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 189642-40-8 REGISTRY
CN **RecA protein (Mycobacterium smegmatis strain mc2-155 clone pRISM1 gene recA) (9CI)** (CA INDEX NAME)
OTHER NAMES:
CN GenBank X99208-derived protein GI 1430892
FS PROTEIN SEQUENCE
MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS, TOXLIT

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 40 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 188900-76-7 REGISTRY
CN **RecA protein (Clostridium perfringens clone pJIR1197 gene recA reduced) (9CI)** (CA INDEX NAME)
OTHER NAMES:
CN GenBank U61497-derived protein GI 1698591
FS PROTEIN SEQUENCE
MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS, TOXLIT

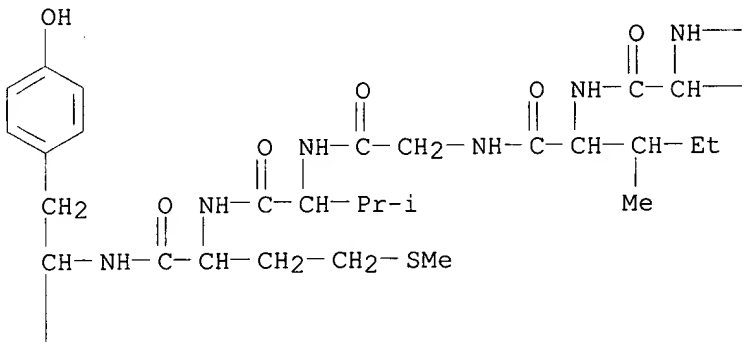
*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
1 REFERENCES IN FILE CA (1967 TO DATE)

1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

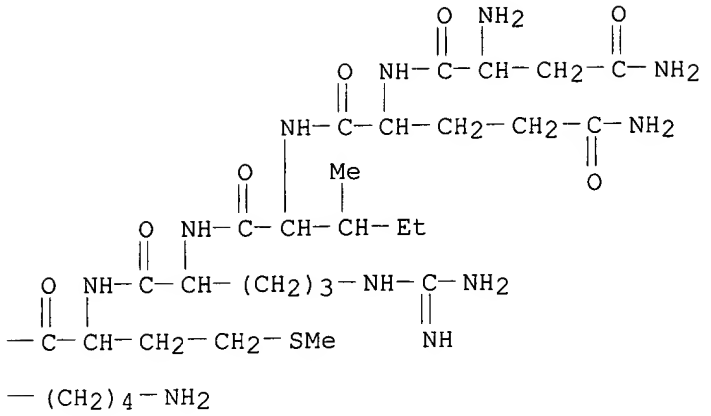
L11 ANSWER 41 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 187484-59-9 REGISTRY
CN **RecA protein (Paracoccus denitrificans clone pUA617 gene recA reduced) (9CI)** (CA INDEX NAME)
OTHER NAMES:
CN GenBank U59631-derived protein GI 1825468
FS PROTEIN SEQUENCE
MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS, TOXLIT

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

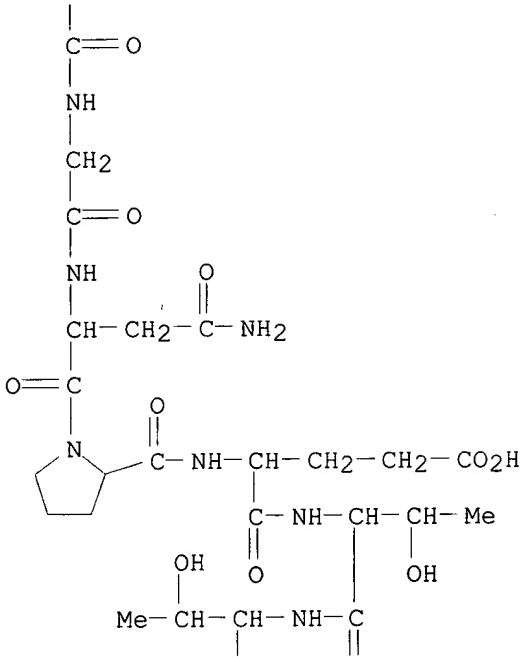
L11 ANSWER 42 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 186422-52-6 REGISTRY
CN Glycine, L-asparaginyl-L-glutaminyl-L-isoleucyl-L-arginyl-L-methionyl-L-lysyl-L-isoleucylglycyl-L-valyl-L-methionyl-L-tyrosylglycyl-L-asparaginyl-L-prolyl-L-.alpha.-glutamyl-L-threonyl-L-threonyl-L-threonylglycyl- (9CI) (CA INDEX NAME)
OTHER NAMES:
CN **193-212-RecA protein peptide YECO (Escherichia coli DNA-binding loop L2)**
FS PROTEIN SEQUENCE; STEREOSEARCH
MF C91 H151 N27 O30 S2
SR CA
LC STN Files: CA, CAPLUS, TOXLIT, USPATFULL



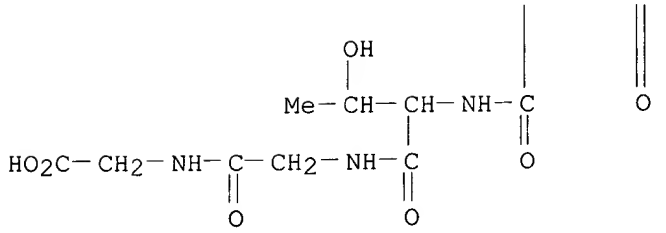
PAGE 1-B



PAGE 2-A



PAGE 3-A



3 REFERENCES IN FILE CA (1967 TO DATE)
3 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 43 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 184787-08-4 REGISTRY
CN **RecA protein (Spiroplasma citri N-terminal fragment) (9CI)** (CA

INDEX NAME)
OTHER NAMES:
CN GenBank U31909-derived protein GI 1199583
FS PROTEIN SEQUENCE
MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS, TOXLIT

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 44 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 184787-07-3 REGISTRY
CN **RecA protein (Spiroplasma melliferum strain BC3 N-terminal fragment) (9CI)** (CA INDEX NAME)
OTHER NAMES:
CN GenBank U43710-derived protein GI 1236436
FS PROTEIN SEQUENCE
MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS, TOXLIT

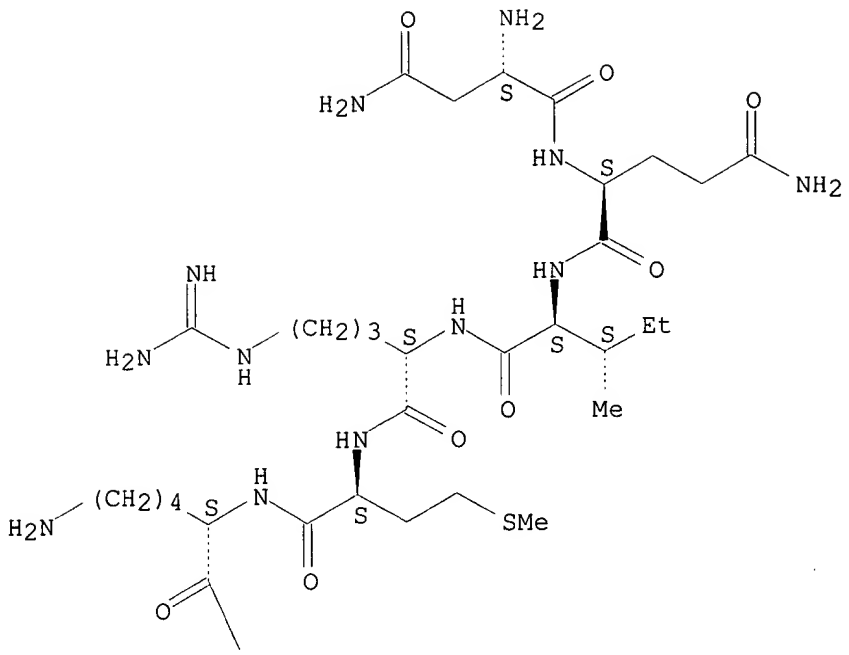
*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

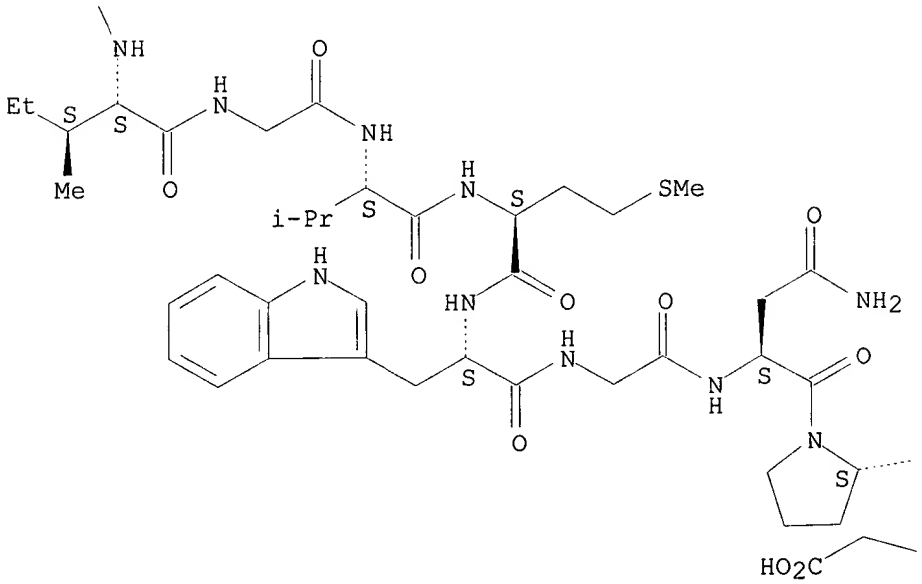
L11 ANSWER 45 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 184693-76-3 REGISTRY
CN **RecA protein (Mycoplasma pneumoniae strain M129 gene recA) (9CI)** (CA INDEX NAME)
OTHER NAMES:
CN GenBank AE000033-derived protein GI 1674029
FS PROTEIN SEQUENCE
MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

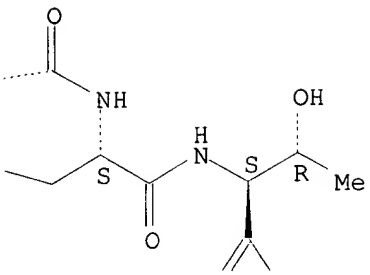
L11 ANSWER 46 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 177158-15-5 REGISTRY
CN Glycine, L-asparaginyl-L-glutaminyl-L-isoleucyl-L-arginyl-L-methionyl-L-lysyl-L-isoleucylglycyl-L-valyl-L-methionyl-L-tryptophylglycyl-L-asparaginyl-L-prolyl-L-.alpha.-glutamyl-L-threonyl-L-threonyl-L-threonylglycyl- (9CI) (CA INDEX NAME)
OTHER NAMES:
CN **193-212-RecA protein peptide WECO (Escherichia coli DNA-binding loop L2)**
CN Peptide WECO (synthetic 20-amino acid)
FS PROTEIN SEQUENCE; STEREOSEARCH
MF C93 H152 N28 O29 S2
SR CA
LC STN Files: CA, CAPLUS, TOXLIT, USPATFULL

Absolute stereochemistry.

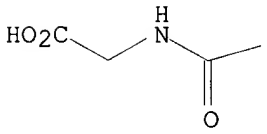




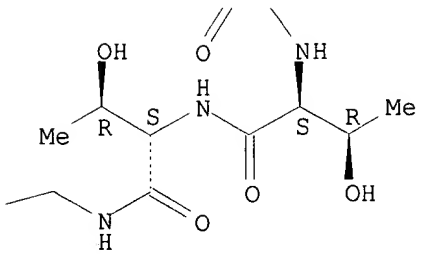
PAGE 2-B



PAGE 3-A



PAGE 3-B



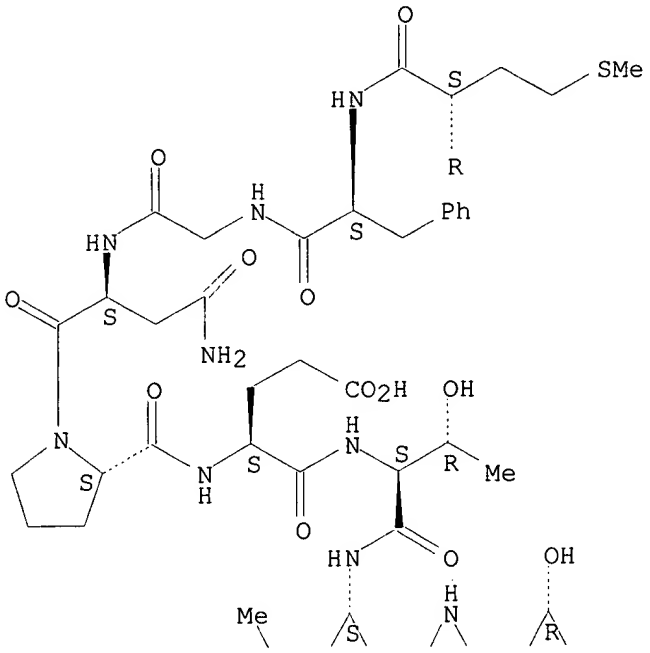
5 REFERENCES IN FILE CA (1967 TO DATE)
5 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 47 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 177158-14-4 REGISTRY
CN Glycine, L-asparaginyl-L-glutaminyl-L-isoleucyl-L-arginyl-L-methionyl-L-lysyl-L-isoleucylglycyl-L-valyl-L-methionyl-L-phenylalanylglycyl-L-asparaginyl-L-prolyl-L-.alpha.-glutamyl-L-threonyl-L-threonyl-L-threonylglycyl- (9CI) (CA INDEX NAME)
OTHER NAMES:

CN **193-212-RecA protein peptide FECO (Escherichia coli DNA-binding loop L2)**
CN Peptide FECO (synthetic 20-amino acid)
FS PROTEIN SEQUENCE; STEREOSEARCH
MF C91 H151 N27 O29 S2
SR CA
LC STN Files: CA, CAPLUS, TOXLIT, USPATFULL

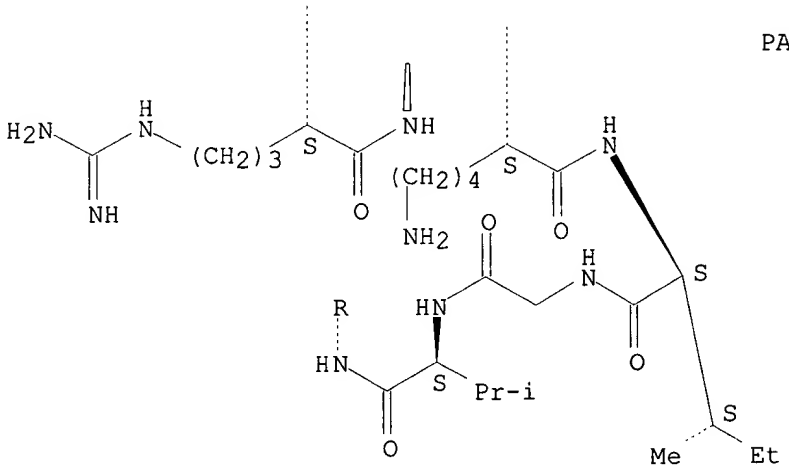
Absolute stereochemistry.

PAGE 1-A



* STRUCTURE DIAGRAM TOO LARGE FOR DISPLAY - AVAILABLE VIA OFFLINE PRINT *

PAGE 3-A



4 REFERENCES IN FILE CA (1967 TO DATE)
4 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 48 OF 53 REGISTRY COPYRIGHT 2000 ACS

RN 168183-38-8 REGISTRY
CN Enzyme (Helicobacter pylori clone pSAT105 gene recA) (9CI) (CA INDEX NAME)
OTHER CA INDEX NAMES:
CN Enzyme (Campylobacter pyloridis clone pSAT105 gene recA)
OTHER NAMES:
CN GenBank U13756-derived protein
CN **RecA protein (Helicobacter pylori clone pSAT105 gene recA)**
FS PROTEIN SEQUENCE
MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS, TOXLIT, USPATFULL

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
2 REFERENCES IN FILE CA (1967 TO DATE)
2 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 49 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 156348-42-4 REGISTRY
CN Enzyme (Rickettsia prowazeki strain Madrid E gene recA reduced) (9CI) (CA INDEX NAME)
OTHER CA INDEX NAMES:
CN Enzyme (Rickettsia prowazekii strain Madrid E gene recA reduced)
OTHER NAMES:
CN GenBank AJ235273-derived protein GI 3861290
CN Protein (Rickettsia prowazekii gene recA)
CN **RecA protein (recA) (Rickettsia prowazeki gene RP761)**
FS PROTEIN SEQUENCE
MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS, TOXLIT

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
2 REFERENCES IN FILE CA (1967 TO DATE)
2 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 50 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 155979-15-0 REGISTRY
CN Enzyme (Campylobacter jejuni clone pMP100/pMP101 gene recA reduced) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN **RecA protein (Campylobacter jejuni strain 81-176 clone pMP100/pMP101 gene recA)**
FS PROTEIN SEQUENCE
MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS, CHEMCATS, CSCHEM

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 51 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 155360-01-3 REGISTRY
CN Enzyme (Mycobacterium leprae clone pEJ217 gene recA reduced) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN GenBank U00019-derived protein GI 467094
CN Protein (Mycobacterium leprae clone pEJ217 gene recA)
CN **RecA protein (Mycobacterium leprae clone B2235 gene recA)**
FS PROTEIN SEQUENCE
MF Unspecified
CI MAN

SR CA
LC STN Files: CA, CAPLUS, TOXLIT

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
3 REFERENCES IN FILE CA (1967 TO DATE)
3 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 52 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 111645-21-7 REGISTRY
CN Protein (Pseudomonas aeruginosa clone pMY21 gene recA reduced) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN GenBank AE004782-derived protein GI 9949775
CN **RecA protein (Pseudomonas aeruginosa strain PAO1 gene recA)**
FS PROTEIN SEQUENCE
MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS, TOXLIT

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
2 REFERENCES IN FILE CA (1967 TO DATE)
2 REFERENCES IN FILE CAPLUS (1967 TO DATE)

L11 ANSWER 53 OF 53 REGISTRY COPYRIGHT 2000 ACS
RN 73177-10-3 REGISTRY
CN Enzyme (Escherichia coli gene recA reduced) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN GenBank AE000354-derived protein GI 1789051
CN **RecA protein (Escherichia coli strain K12-MG1655 gene recA)**
FS PROTEIN SEQUENCE
MF Unspecified
CI MAN
LC STN Files: BIOTECHNO, CA, CAPLUS, EMBASE, TOXLIT

*** STRUCTURE DIAGRAM IS NOT AVAILABLE ***
*** USE 'SQD' OR 'SQIDE' FORMATS TO DISPLAY SEQUENCE ***
4 REFERENCES IN FILE CA (1967 TO DATE)
4 REFERENCES IN FILE CAPLUS (1967 TO DATE)

WEST

Freeform Search

Database: US Patents Full-Text Database
JPO Abstracts Database
EPO Abstracts Database
Derwent World Patents Index
IBM Technical Disclosure Bulletins

Term:
```
reca
e adj
```

Display: 10 **Documents in Display Format:** TI **Starting with Number** 1

Generate: ○ Hit List ◉ Hit Count ○ Image

Search | Clear | Help | Logout | Interrupt

Main Menu | Show S Numbers | Edit S Numbers | Preferences

Search History

Today's Date: 11/30/2000

DB Name	Query	Hit Count	Set Name
USPT	l4 and l3 and l2 and l1	3	L5
USPT	((435/252.33)!.CCLS.)	1869	L4
USPT	((435/195)!.CCLS.)	326	L3
USPT	e adj coli	25526	L2
USPT	reca	968	L1

WEST

Generate Collection

Search Results - Record(s) 1 through 3 of 3 returned.

☐ 1. Document ID: US 5919648 A

L5: Entry 1 of 3 File: USPT Jul 6, 1999

US-PAT-NO: 5919648
DOCUMENT-IDENTIFIER: US 5919648 A

TITLE: Industrial enzymes

DATE-ISSUED: July 6, 1999

INVENTOR-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY
Burr; Keith William	Ulverston	N/A	N/A	GBX
Ramsden; Martin	Ulverston	N/A	N/A	GBX
Illing; Graham Timothy	Montrose	N/A	N/A	GBX
Harrison; Leslie Ann	Ulverston	N/A	N/A	GBX
Maishman; Nicholas John	Beckenham	N/A	N/A	GBX
Spence; David Wilson	Ulverston	N/A	N/A	GBX
Slade; Andrew	Ulverston	N/A	N/A	GBX

ASSIGNEE INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY	TYPE CODE
Glaxo Group Limited	Greenford	N/A	N/A	GBX	03

APPL-NO: 8/ 817900
DATE FILED: April 25, 1997

FOREIGN-APPL-PRIORITY-DATA:

COUNTRY	APPL-NO	APPL-DATE
GB	9423212	November 17, 1994

PCT-DATA:

APPL-NO	DATE-FILED	PUB-NO	PUB-DATE	371-DATE	102(E)-DATE
PCT/EP95/04487	Nov 15, 1995	WO96/16174	May 30, 1996	Apr 25, 1997	Apr 25, 1997

INT-CL: [6] C12P 35/00, C12N 9/14, C12N 1/20, C12N 15/00
US-CL-ISSUED: 435/47; 435/195, 435/252.33, 435/320.1, 435/874, 435/69.1, 536/23.2
US-CL-CURRENT: 435/47; 435/195, 435/252.33, 435/320.1, 435/69.1, 435/874, 536/23.2
FIELD-OF-SEARCH: 435/47, 435/195, 435/69.1, 435/71.2, 435/252.3, 435/252.33, 435/320.1, 435/874, 536/23.2

REF-CITED:

U.S. PATENT DOCUMENTS

PAT-NO	ISSUE-DATE	PATENTEE-NAME	US-CL
5320948	June 1994	Iwami et al.	435/47

ART-UNIT: 162
PRIMARY-EXAMINER: Patterson, Jr.; Charles L.
ASSISTANT-EXAMINER: Saidha; Tekchand
ATTY-AGENT-FIRM: Nixon & Vanderhye P.C.

ABSTRACT:

The present invention relates to an enzyme process for the one-step conversion of cephalosporin C or a derivative thereof into 7-aminocephalosporanic acid or a corresponding derivative thereof. The one step conversion is effected using a cephalosphorin C amidohydrolase derived from Pseudomonas Vesicularis B965, or from any cephalosporin C amidolydrolase producing or potentially producing descendants thereof, or from expression of DNA derived from Pseudomonas Vesicularis B965 or any cephalosporin C amidohydrolase producing or potentially producing descendants thereof

8 Claims, 1 Drawing figures

☐ 2. Document ID: US 5858754 A

L5: Entry 2 of 3 File: USPT Jan 12, 1999

US-PAT-NO: 5858754
DOCUMENT-IDENTIFIER: US 5858754 A

TITLE: Methods of analysis and manipulation of DNA utilizing mismatch repair systems

DATE-ISSUED: January 12, 1999

INVENTOR-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY
Modrich; Paul L.	Chapel Hill	NC	N/A	N/A
Su; Shin-San	Newton	MA	N/A	N/A
Au; Karin G.	Durham	NC	N/A	N/A
Lahue; Robert S.	Northboro	MA	N/A	N/A
Cooper; Deani Lee	Watertown	MA	N/A	N/A
Worth, Jr.; Leroy	Durham	NC	N/A	N/A

ASSIGNEE INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY	TYPE CODE
Duke University	Durham	NC	N/A	N/A	02

APPL-NO: 8/ 600450
DATE FILED: February 13, 1996

PARENT-CASE:
This is a continuation of application Ser. No. 08/145,837, filed Nov. 1, 1993, now U.S. Pat. No. 5,556,750 noted Jun. 30, 1998 hereby incorporated by reference in its totality (including drawings) and is a continuation in part of Ser. No. 08/002,529, filed Jan. 11, 1993 now abandoned which is a continuation of Ser. No. 07/350,983, filed May 12, 1989, now abandoned.

INT-CL: [6] C12N 9/14, C12N 9/06
US-CL-ISSUED: 435/195; 435/6, 435/191, 435/252.3, 435/252.33, 435/91.1, 536/23.1, 536/23.2, 536/23.4, 536/23.7
US-CL-CURRENT: 435/195; 435/191, 435/252.3, 435/252.33, 435/6, 435/91.1,

536/23.1, 536/23.2, 536/23.4, 536/23.7
FIELD-OF-SEARCH: 435/191, 435/6, 435/91.1, 435/91.2, 435/195, 435/252.3, 435/252.33, 536/24.3, 536/24.31, 536/24.32, 536/23.1, 536/23.2, 536/23.4, 536/23.7

REF-CITED:

U.S. PATENT DOCUMENTS

PAT-NO	ISSUE-DATE	PATENTEE-NAME	US-CL
4794075	December 1988	Ford et al.	435/6

FOREIGN PATENT DOCUMENTS

FOREIGN-PAT-NO	PUBN-DATE	COUNTRY
2239456	July 1991	GBX
9302216	February 1993	WOX
9320233	October 1993	WOX
9322457	November 1993	WOX
9322462	November 1993	WOX

OTHER PUBLICATIONS

Dervan, P. B. "Design of Sequence-Specific DNA-Binding Molecules" Science 232, 464-471, Apr. 25, 1986.
Dreyer et al. "Sequence-specific cleavage of Single-Strand DNA: Oligo . . . " Proc. Natl. Acad. Sci. USA 82, 968-972, Feb. 1985.
Adams et al., "The Biochemistry of the Nucleic Acids," Chapman & Hall pp. 221-223 (1986).
Au et al., "Initiation of Methyl-directed Mismatch Repair," J. Biol. Chem. 267:12142-12148 (1992).
Au et al., "Escherichia coli mutY Gene Encodes An Adenine Glycosylase Active on G-A Mispairs," Proc. Natl. Acad. Sci. USA 86:8877-8881 (1989).
Chen and Sigman, "Chemical Conversion of a DNA-Binding Protein Into a Site-Specific Nuclease," Science 237:1197-1201 (1987).
Cotton et al., "Reactivity of Cytosine and Thymine In Single-Base-Pair Mismatches with Hydroxylamine and Osmium Tetroxide and Its Application to the Study of Mutations," Proc. Natl. Acad. Sci. USA 85:4397-4401 (1986).
Grilley et al., "Mechanisms of DNA-Mismatch Correction," Mutation Research 236:253-267 (1990).
Grilley et al., "Isolation and Characterization of the Escherichia coli mutL Gene Product," J. Biol. Chem. 264:1000-1004 (1989).
Hennighausen and Lubon, "Interaction of Protein With DNA In Vitro," Guide to Molecular Cloning Techniques, Berger and Kimmel eds., 152:721-735 (1987).
Holmes et al., "Strand-specific Mismatch Correction In Nuclear Extracts of Human and Drosophila Melanogaster Cell Lines," Proc. Natl. Acad. Sci. USA 87:5837-5841 (1990).
Jiricny et al., "Mistmatch-containing Oligonucleotide Duplexes Bound By the E. coli mutS-encoded Protein," Nucleic Acids Research 16:7843-7853 (1988).
Lahue and Modrich, "Methyl-directed DNA Mismatch Repair in Eschericha coli," Mutation Research 198:37-43 (1988).
Lahue et al., "Requirement for d(GATC) Sequences in Escherichia coli mutHLS Mismatch Correctiun," Proc. Natl. Acad. Sci. USA 84:1482-1486 (1987).
Lahue and Modrich, "DNA Mismatch Correction in a Defined System," Science 245:160-164 (1989).
Lu and Hsu, "Detection of Single DNA Based Mutations with Mismatch Repair Enzymes," Genomics 14:249-255 (1992).
Lu and Chang, "A Novel Nucelotide Excision Repair for The Conversion of An A/G Mismatch to C/G Base Pair in E. coli," Cell 54:805-812 (1988).
Lu et al., "Repair of DNA Base-pair Mismatches in Extracts of Escherichia coli," Cold Spring Harbor Laboratory, Cold Spring Harbor Symposia on Quantitative Biology, XLIX:589-596 (1984).
Marx "DNA Repair Comes Into Its Own," Science 266:728-730 (1994).
Modrich et al., "DNA Mismatch Correction," Ann. Rev. Biochem. 56:435-466 (1987).

Modrich, "Methyl-directed DNA Mismatch Correction," J. Biol. Chem. 264:6597-6600 (1989).
Modrich, "Mismatch Repair, Genetic Stability and Cancer," Science 266:1959-1960 (1994).
Modrich, Molecular Mechanisms of DNA--Protein Interaction, 1986, NIH Grant, Abstract (Source: CRISP).
Modrich, Molecular Mechanisms of DNA--Protein Interaction, 1987, NIH Grant, Abstract (Source: CRISP).
Modrich, Molecular Mechanisms of DNA--Protein Interaction, 1988, NIH Grant, Abstract (Source: CRISP).
Modrich, Molecular Mechanisms of DNA--Protein Interaction, 1989, NIH Grant, Abstract (Source: CRISP).
Modrich, Molecular Mechanisms of DNA--Protein Interaction, 1990, NIH Grant, Abstract (Source: CRISP).
Modrich, Molecular Mechanisms of DNA--Protein Interaction, 1991, NIH Grant, Abstract (Source: CRISP).
Modrich, Molecular Mechanisms of DNA--Protein Interaction, 1992, NIH Grant, Abstract (Source: CRISP).
Modrich, Molecular Mechanisms of DNA--Protein Interaction, 1993, NIH Grant, Abstract (Source: CRISP).
Modrich, Enzymology of Eukaryotic DNA Mismatch Repair, 1991, NIH Grant, Abstract (Source: CRISP).
Modrich, Enzymology of Eukaryotic DNA Mismatch Repair, 1992, NIH Grant, Abstract (Source: CRISP).
Modrich, Enzymology of Eukaryotic DNA Mismatch Repair, 1993, NIH Grant, Abstract (Source: CRISP).
Priebe et al., "Nucleotide Sequence of the hexA Gene for DNA Mistmatch Repair in Streptococus pneumoniae and Homology of hexA to mutS of Escherichia coli and Salmonella typhimurium," J. Bacteriology 170:190-196 (1988).
Quinones et al., "Expression of the Escherichia coli dna Q (mutD) Gene is Inducible," Mol. Gene Genet. 211:106-112 (1988).
Su and Modrich, "Escherichia coli mutS-encoded Protein Binds to Mismatched DNA Base Pairs," Proc. Natl. Acad. Sci. USA 83:5057-5061 (1986).
Su et al., "Mispair Specificity of Methyl-directed DNA Mismatch Correction in Vitro," J. Biol. Chem. 263:6829-6835 (1988).
Su et al., "Gap Formation is Associated With Methyl-Directed Mismatch Correction Under Conditions of Restricted DNA Synthesis," Genome 31:104-111 (1989).
Welsh et al., "Isolation and Characterization of the Escherichia coli mush Gene Product," J. Biol. Chem. 262:15624-15629 (1987).
Wilchele et al., Analytical Biochem. 171:1-32 (1988).

ART-UNIT: 162
PRIMARY-EXAMINER: Wax; Robert A.
ASSISTANT-EXAMINER: Nashed; Nashaat T.
ATTY-AGENT-FIRM: Lyon & Lyon LLP

ABSTRACT:

A diagnostic method for detecting a base pair mismatch in a DNA duplex, comprising the steps of contacting at least one strand of a first DNA molecule with the complementary strand of a second DNA molecule under conditions such that base pairing occurs contacting a DNA duplex potentially containing a base pair mismatch with a mispair recognition protein under conditions suitable for the protein to form a specific complex only with the DNA duplex having a base pair mismatch, and not with a DNA duplex lacking a base pair mismatch, and detecting any complex as a measure of the presence of a base pair mismatch in the DNA duplex.

5 Claims, 10 Drawing figures

☐ 3. Document ID: US 5589375 A

L5: Entry 3 of 3 File: USPT Dec 31, 1996

US-PAT-NO: 5589375
DOCUMENT-IDENTIFIER: US 5589375 A

TITLE: PTP 1D: a novel protein tyrosine phosphatase

DATE-ISSUED: December 31, 1996

INVENTOR-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY
Ullrich; Axel	Martinsried bei Muchen	N/A	N/A	DEX
Vogel; Wolfgang	Germering	N/A	N/A	DEX

ASSIGNEE INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY	TYPE CODE
Max-Planck-Gesellschaft Zur Forderung Der Wissenschaften E. V.	Gottingen	N/A	N/A	DEX	03

APPL-NO: 8/ 018129
DATE FILED: February 16, 1993

PARENT-CASE:
1. CROSS REFERENCE TO RELATED APPLICATION This application is a continuation-in-part of U.S. application Ser. No. 07/956,315, filed Oct. 6, 1992, now abandoned which is incorporated herein by reference in its entirety.

INT-CL: [6] C12N 5/06, C07H 21/04
US-CL-ISSUED: 435/240.2; 435/172.3, 435/320.1, 435/252.3, 435/252.33, 435/4, 435/195, 935/14, 935/20, 935/24, 935/32, 935/70, 935/71, 935/72, 935/73, 536/23.2, 536/23.1
US-CL-CURRENT: 435/325; 435/195, 435/252.3, 435/252.33, 435/320.1, 435/357, 435/358, 435/365, 435/369, 435/64, 536/23.1 , 536/23.2
FIELD-OF-SEARCH: 536/23.2, 536/23.1, 435/194, 435/195, 435/320.1, 435/91.1, 435/91.2, 435/240.2, 435/252.3, 435/252.33, 435/172.3, 930/240, 935/14, 935/29, 935/24, 935/32, 935/70, 935/71, 935/72, 935/73

REF-CITED:

U.S. PATENT DOCUMENTS

PAT-NO	ISSUE-DATE	PATENTEE-NAME	US-CL
4675285	June 1987	Clark et al.	435/6

FOREIGN PATENT DOCUMENTS

FOREIGN-PAT-NO	PUBN-DATE	COUNTRY
9201050	January 1992	WOX

OTHER PUBLICATIONS

Freeman et al. 1992, Proc. Natl. Acad. Sci., USA., 89, 11239-11243.
Mathews et al. 1992. Mol. Cell. Biol. 12, 2396-2405.
Plutzky et al. 1992. Proc. Natl. Acad. Sci. USA 89, 1123-1127.
Yi, et al 1992. Mol. Cell. Biol. 12, 836-846.
Gu et al. 1991 Proc. Natl Acad Sci. USA. 88, 5867-5871.
Lombroso et al 1991 Proc. Natl. Acad. Sci USA 88, 7242-7246.
Shen et al. 1991 Nature 352, 736-739.
Yang et al. 1991. Proc. Natl Acad Sci USA 88, 5949-5953.
Chernoff et al. 1990 Proc. Natl. Acad. Sci. USA. 87, 2735-2739.
Guan et al, 1990. Proc. Natl. Acad. Sci. USA. 87, 1501-1505.
Charbonneau et al. 1989. Proc. Natl Acad Sci USA 86, 5252-5256.

Cool et al. 1989. Proc. Natl Acad Sci USA 86, 5257-5261.
Pallen et al. 1988 Ann. N.Y. Acad Sci. 551, 299-308.
Tonks et al. 1988 J. Biol. Chem. 263, 6722-6730.
Gebbink et al. 1991. FEBS Lett. 290, 123-130.
Jirik et al. 1990 FEBS Lett. 273, 239-242.
Kaplan et al. 1990. Proc. Natl Acad Sci USA 87, 7000-7004.
Krueger et al. 1990 EMBO J. 9, 3241-3252.
Mathews et al. 1990 Proc. Natl. Acad. Sci USA 87, 4444-4448.
Mustelin et al. 1989 Proc. Natl. Acad Sci USA 86, 6302-6306.
Streuli et al. 1988. J. Exp. Med. 168, 1523-1530.
Ralph et al. 1987, EMBO. J. 6, 1251-1257.
Perkins et al. 1992 Cell. 70, 225-236.
Harihan et al. 1991 Proc. Natl. Acad. Sci. USA. 88, 11266-11270.
Streuli et al. 1989 Proc. Natl. Acad Sci. USA. 86, 8698-8702.
Sprenger et al. 1989. Nature 338, 478-483.
Russell et al. 1992 F.E.B.S. 304, 15-20.
Zheng et al. 1992 Nature 359, 336-339.
Stover et al. 1991. Proc. Natl. Acad. Sci. USA 88, 7704-7707.
Margolis et al. 1989 Cell 57, 1101-1107.
Morla et al. 1989. Cell 58, 193-203.
Charbonneau et al, 1992. Annu. Rev. Cell Biol. 8, 463-493.
Pot et al. 1992 Biochim. Biophys. Acta. 1136, 35-43.
Fischer et al. Scierre 253, 401-406.
Koch et al. 1991 Science 252, 668-674.
Hunter, T. 1989. Cell 58, 1013-1016.
Hunter, T. 1987. Cell, 49, 1-4.
Watson, J. D. 1987, in:Molecular Biology of The Gene., 3rd Edition.
Benja/Cummings Publ. Co., Inc. Menlo Park, CA 94025. p. 313.
Hackh's Chemical Dictionary, 1969. (Grant, J., ed.) McGrawHill Book Co., NY, p. 236.
Vogel et al. 1993. Science 259, 1611-1614.
Baker, et al., 1982, in: The Study of Biology. Fourth edition. Addison-Wooley Publ. Co., Reading, MA. p. 9.
Co et al, 1992. Biophys J. 61 (2 part 2), A337, abstract 1940.
Day, R. A. 1983. How to Write and Publish a scientific Paper. Second Edition. ISI Press. Philadelphia, PA pp. 15-19.
Zhang et al. 1991 Biochem. Biophys. Res. Commun 178, 1291-1297.

ART-UNIT: 184
PRIMARY-EXAMINER: Low; Christopher S.
ATTY-AGENT-FIRM: Pennie & Edmonds

ABSTRACT:

A novel protein tyrosine phosphatase is the protein designated PTP 1D. The PTP 1D protein may be produced by recombinant means, for example using a nucleic acid construct encoding the protein as provided herein. Also disclosed is an antibody specific for an epitope of PTP 1D, protein. Methods for identifying compounds which bind to a PTP 1D protein and inhibit or stimulate its enzymatic activity, pharmaceutical compositions comprising PTP 1D, and methods for treating a disease associated with PTP 1D protein using such compositions, are provided.

10 Claims, 15 Drawing figures

Display 10 Documents, starting with Document: 3

Display Format: FRO Change Format

Search Results -

Terms	Documents
bind and l11	221

Database: US Patents Full-Text Database
JPO Abstracts Database
EPO Abstracts Database
Derwent World Patents Index
IBM Technical Disclosure Bulletins

Refine Search: bind and l11 Clear

Search History

Today's Date: 11/30/2000

DB Name	Query	Hit Count	Set Name
USPT	bind and l11	221	L14
USPT	l12 and l1	0	L13
USPT	MAW	251	L12
USPT	l8 and l4	320	L11
USPT	l9 and l4	231	L10
USPT	l5 and l6	442	L9
USPT	l5 and l7	622	L8
USPT	l1 and l3	876	L7
USPT	l1 and l2	567	L6
USPT	mutant	19936	L5
USPT	ATP	12341	L4
USPT	e adj coli	25526	L3
USPT	escherichia adj coli	16935	L2
USPT	recA	968	L1

L16: Entry 1 of 107 File: USPT Nov 7, 2000

DOCUMENT-IDENTIFIER: US 6143566 A
TITLE: Methods of performing homologous recombination based modification of nucleic acids in recombination deficient cells and use of the modified nucleic acid products thereof

BSPR:
Functional analyses of genes in vivo frequently involve the introduction of modified genomic DNA into the germline to generate transgenic animals [Jaenisch et al., Science 240:1468 (1985); Brinster, Cell 41:343 (1985)]. The genomic DNA sequences containing introns and essential regulatory sequences have been shown to be expressed in vivo in cases where simple cDNA constructs cannot be expressed [Brinster et al., Proc. Natl. Acad. Sci. 85:836-840 (1988)]. Furthermore, the size of the genomic DNA that can be readily manipulated in vitro and introduced into the germline can be a critical determinant of the outcome of the functional analysis of a gene since elements that are important for high level, tissue specific and position-independent expression of the transgene may be located at a long distance from the gene itself [Dillon et al., Trends Genet. 9:134 (1993); Kennison, Trends Genet. 9:75 (1993); Wilson et al., Annu. Rev. Cell. Biol. 6:679 (1990)].

BSPR:
On the other hand, the use of such large genomic transgenes has several practical problems. For example, the size of the transgene is presently limited due to constraints on the sequence length that can be cloned and stably maintained in a conventional plasmid or a cosmid. Thus DNA sequences suspected of being nonessential are often omitted when designing the constructs to be transferred because of the size limitation. In addition, in vitro manipulations of large DNAs oftentimes lead to mechanical shear [Peterson et al., TIG 13:61-66].

BSPR:
Yeast artificial chromosomes (YACs) allow large genomic DNA to be modified and used for generating transgenic animals [Burke et al., Science 236:806; Peterson et al., Trends Genet. 13:61 (1997); Choi, et al., Nat. Genet., 4:117-223 (1993), Davies, et al., Biotechnology 11:911-914 (1993), Matsuura, et al., Hum. Mol. Genet., 5:451-459 (1996), Peterson et al., Proc. Natl. Acad. Sci., 93:6605-6609 (1996); and Schedl, et al., Cell, 86:71-82 (1996)]. Other vectors also have been developed for the cloning of large segments of mammalian DNA, including cosmids, and bacteriophage P1 [Sternberg et al., Proc. Natl. Acad. Sci. U.S.A., 87:103-107 (1990)]. YACs have certain advantages over these alternative large capacity cloning vectors [Burke et al., Science, 236:806-812 (1987)]. The maximum insert size is 35-30 kb for cosmids, and 100 kb for bacteriophage P1, both of which are much smaller than the maximal insert for a YAC. However, there are several critical limitations in the YAC system including difficulties in manipulating YAC DNA, chimerism and clonal instability [Green et al., Genomics, 11:658 (1991); Kouprina et al., Genomics 21:7 (1994); Larionov et al., Nature Genet. 6:84 (1994)]. As a result, generating transgenic mice with an intact YAC remains a challenging task [Burke et al., Science 236:806; Peterson et al., Trends Genet. 13:61 (1997)].

BSPR:
An alternative to YACs are E. coli based cloning systems based on the E. coli fertility factor that have been developed to construct large genomic DNA insert libraries. They are bacterial artificial chromosomes (BACs) and P-1 derived artificial chromosomes (PACs) [Mejia et al., Genome Res. 7:179-186 (1997); Shizuya et al., Proc. Natl. Acad. Sci. 89:8794-8797 (1992); Ioannou et al., Nat. Genet., 6:84-89 (1994); Hosoda et al., Nucleic Acids Res. 18:3863 (1990)]. BACs are based on the E. coli fertility plasmid (F factor); and PACs are based on the bacteriophage P1. The size of DNA fragments from eukaryotic genomes that can be stably cloned in Escherichia coli as plasmid molecules has been expanded by the advent of PACs and BACs. These vectors propagate at a very low copy number (1-2 per cell) enabling genomic inserts up to 300 kb in size to be stably maintained in recombination deficient hosts (most clones in human genomic libraries fall within the 100-200 kb size range). The host cell is required to be recombination deficient to ensure that

non-specific and potentially deleterious recombination events are kept to a very minimum. As a result, libraries of PACs and BACs are relatively free of the high proportion of chimeric or rearranged clones typical in YAC libraries, [Monaco et al., Trends Biotechnol 12:280-286 (1994); Boyseu et al., Genome Research, 7:330-338 (1997)]. In addition, isolating and sequencing DNA from PACs or BACs involves simpler procedures than for YACs, and PACs and BACs have a higher cloning efficiency than YACs [Shizuya et al., Proc. Natl. Acad. Sci. 89:8794-8797 (1992); Ioannou et al., Nat. Genet., 6:84-89 (1994); Hosoda et al., Nucleic Acids Res. 18:3863 (1990)]. Such advantages have made BACs and PACs important tools for physical mapping in many genomes [Woo et al., Nucleic Acids Res., 22:4922 (1994); Kim et al., Proc. Natl. Acad. Sci. 93:6297-6301 (1996); Wang et al., Genomics 24:527 (1994); Wooster et al., Nature 378:789 (1995)]. Furthermore, the PACs and BACs are circular DNA molecules that are readily isolated from the host genomic background by classical alkaline lysis [Birnboim et al., Nucleic Acids Res. 7:1513-1523 (1979].

BSPR:
Functional characterization of a gene of interest contained by a PAC or BAC clone generally entails transferring the DNA into a eukaryotic cell for transient or long-term expression. A transfection reporter gene, e.g., a gene encoding lacZ, together with a selectable marker, e.g., neo, can be inserted into a BAC [Mejia et al., Genoine Res. 7:179-186 (1997). Transfected cells can be then detected by staining for X-Gal to verify DNA uptake. Stably transformed cells are selected for by the antibiotic G418.

BSPR:
However, while PACs and BACs have cloning capacities up to 350 kb, performing homologous recombination to introduce mutations into a gene of interest has not been demonstrated [Peterson et al., TIG 13:61-66]. Indeed, although BACs or PACs have become an important source of large genomic DNA in genome research, there are still no methods available to modify the BACs or PACs. Furthermore, no germline transmission of intact BACs or PACs in transgenic mice have been reported. These, as well as other disadvantages of BACs and PACs greatly limit their potential use for functional studies. Therefore, there is a need for an improved cloning vector for germline transmission of selected genes in transgenic animals. More particularly there is a need for a cloning vector that has the capacity to contain greater than 100 kilobases of DNA, which can be readily manipulated and isolated, but still can be stably stored in libraries relatively free of rearranged clones. In addition, there is a need to provide methodology for generating such cloning vectors. There is also a need to apply such vectors to improve current technologies such as gene targeting.

BSPR:
Gene targeting has been used in various systems, from yeast to mice, to make site specific mutations in the genome. Gene targeting is not only useful for studying function of proteins in vivo, but it is also useful for creating animal models for human diseases, and in gene therapy. The technique involves the homologous recombination between DNA introduced into a cell and the endogenous chromosomal DNA of the cell. However, in the vertebrate system, the rate of homologous recombination is very low, as compared to random integration. The only cell line that allows a relatively high homologous recombination rate and maintains the ability to populate the germline is the murine 129 embryonic stem cells (ES cells). Using this specialized cell, mice can be generated with a targeted mutation (Gene Targeting, a practical approach Ed. by A. Joyner, IRL Press: Oxford, New York, Tokyo). However, the rate of homologous recombination for some gene loci in ES cells is still extremely low (<1%), the procedure is labor intensive, and the cost of generating targeted mutant mice is very expensive. Moreover, since there are no ES cells available for vertebrates other than mice, gene targeting in a germline is still not possible for other vertebrates.

BSPR:
The major limitation for gene targeting in vertebrate cells remain to be the low targeting frequency. One critical factor affecting the targeting frequency is the total length of homology. Deng and Capecchi (MCB, 12:3365-3371) have shown that gene targeting frequency is linearly-dependent on the logarithm of the total homology length over homology lengths of 2.8 kb to 14.6 kb. Since the curve did not plateau at the 14.6 kb homology, it is likely that incorporating greater homology lengths into the targeting vector will further increase the homologous recombination rate. Using a mathematical model developed by Fujitani et al, [Genetics, 140:797-809, (1995)], an estimate can be made that with a total homology of 100 kb isogenous DNA (i.e., DNA from the same strain of mice), the gene targeting rate in ES cells would be 10%. This is a dramatic improvement over the conventional 14.6 kb targeting vector, which only yields a corresponding rate of only 0.03%. Further support for the present strategy i.e., using a large DNA construct for gene targeting rate comes from an experiment with Mycobacterium tuberculosis, the causal agent of tuberculosis. Like vertebrate

cells, gene targeting in TB has a very low rate, mainly due to the predominance of random integration over homologous recombination. It has been demonstrated that using a 40-50 kb linear targeting construct, a 6% targeting frequency could be obtained, whereas no targeting event was obtained at all with a smaller (<10 kb) targeting construct [Balasubramanian et al., J. of Bacteriology 178:273-279 (1996)]. Therefore, there is a need to construct large gene targeting constructs to allow efficient gene targeting in many biological systems.

BSPR:
In a particular aspect of the present invention, the recombination deficient host cell cannot independently support homologous recombination because the host cell is RecA.sup.-. In this aspect of the invention, inducing the host cell to transiently support homologous recombination comprises inducing the transient expression of a RecA-like protein in the host cell. In a preferred embodiment, inducing the transient expression of the RecA-like protein can be performed with a conditional replication shuttle vector. In a more preferred embodiment the conditional replication shuttle vector is a temperature sensitive shuttle vector (TSSV) that replicates at a permissive temperature, but does not replicate at a non-permissive temperature.

BSPR:
In one particular embodiment of this type, inducing the transient expression of the RecA-like protein comprises transforming the host cell with the TSSV at a permissive temperature, and growing the host cell at a non-permissive temperature. The TSSV encodes a RecA-like protein that is expressed in the host cell and supports the homologous recombination between a nucleic acid contained in a recombination cassette and the particular nucleotide sequence contained in the host cell. The TSSV encoding the RecA-like protein is diluted out when the host cell is grown at the non-permissive temperature. In one particular embodiment of this type the permissive temperature is 30.degree. C. and the non-permissive temperature is 43.degree. C.

BSPR:
In a more intricate version of the present invention, the particular nucleotide sequence which has been selected to undergo homologous recombination is contained in an independent origin based cloning vector (IOBCV) that is comprised by the host cell, and neither the independent origin based cloning vector alone, nor the independent origin based cloning vector in combination with the host cell, can independently support homologous recombination. In a particular embodiment of this type both the independent origin based cloning vector and the host cell are RecA.sup.-, and inducing the host cell to transiently support homologous recombination comprises inducing the transient expression of the RecA-like protein to support homologous recombination in the host cell. In one particular embodiment the independent origin based cloning vector is a Bacterial or Bacteriophage-Derived Artificial Chromosome (BBPAC) and the host cell is a host bacterium.

BSPR:
In a preferred embodiment, inducing the transient expression of the RecA-like protein is performed with a conditional replication shuttle vector that encodes the RecA-like protein. In a more preferred embodiment the conditional replication shuttle vector is a temperature sensitive shuttle vector (TSSV) that replicates at a permissive temperature, but does not replicate at a non-permissive temperature. In one particular embodiment of this type the permissive temperature is 30.degree. C. and the non-permissive temperature is 43.degree. C.

BSPR:
In one embodiment the RecA-like protein is controlled by an inducible promoter and the transient expression of the RecA-like protein is achieved by the transient induction of the inducible promoter in the host cell. In another embodiment, the RecA-like protein is controlled by a constitutive promoter with the transient expression induced by the TSSV.

BSPR:
In a preferred embodiment the TSSV also comprises a recombination cassette and a first gene which bestows resistance to a host cell that contains the TSSV against a first toxic agent. In addition, the first gene can be counter-selected against. The recombination cassette, the RecA-like protein gene, and the first gene are linked together on the TSSV such that when the nucleic acid integrates (i.e. resolved) into the particular nucleotide sequence, the RecA-like protein gene and the first gene remain linked together, and neither the RecA-like protein gene nor the first gene remain linked to the integrated nucleic acid.

BSPR:
In a particular embodiment of this type, the independent origin based cloning vector is a BBPAC and the host cell is a bacterium. The BBPAC further contains a second gene

that bestows resistance to the host cells against a second toxic agent. Introducing the recombination cassette into the host cells is performed by transforming the host cell with the TSSV. Inducing the transient expression of the RecA-like protein to support homologous recombination comprises: (i) incubating the host cells at a permissive temperature in the presence of the first toxic agent and the second toxic agent, wherein transformed host cells containing the TSSV and the BBPAC are selected for and wherein the RecA-like protein is expressed. A first homologous recombination event occurs between the recombination cassette and the particular nucleotide sequence forming a co-integrate between the TSSV and the BBPAC, wherein the TSSV is either free or part of a co-integrate; (ii) incubating the transformed host cells at a non-permissive temperature in the presence of the first toxic agent and the second toxic agent, wherein host cells containing a TSSV co-integrate are selected for, and wherein free TSSV cannot replicate; (iii) selecting a host cell containing a co-integrate between the TSSV and the BBPAC by Southern analysis; (iv) incubating the host cells containing a co-integrate between the TSSV and the BBPAC at a non-permissive temperature in the presence of the second toxic agent, wherein a second homologous recombination event occurs between the recombination cassette and the particular nucleotide sequence, therein integrating the nucleic acid into the particular nucleotide sequence and forming a resolved host cell, i.e., a host cell containing a resolved BBPAC; and (v) incubating the host cells containing the resolved BBPAC in the presence of the second toxic agent, and a counter-selecting agent, and wherein the counter-selecting agent is toxic to host cells containing the first gene, and wherein host cells containing the RecA-like protein gene are removed. Another embodiment further comprises selecting a host cell containing the resolved BBPAC by colony hybridization with a labeled probe that binds to a DNA homologue of the nucleic acid, an mRNA homologue of the nucleic acid, and/or a protein encoded by the nucleic acid. In a particular embodiment, the permissive temperature is 30.degree. C., the non-permissive temperature is 43.degree. C. In a preferred embodiment the incubating of host cells containing the resolved BBPAC in the presence of the second toxic agent and counter-selecting agent is performed at 37.degree. C.

BSPR:
In a particular embodiment the first gene confers tetracycline resistance and the counter-selecting agent is fusaric acid. In a preferred embodiment the RecA-like protein is recA. In the more preferred embodiment the TSSV is pSV1.RecA having the ATCC no. 97968.

BSPR:
In a related aspect of the present invention the RecA-like protein is controlled by an inducible promoter, and the transient expression of the RecA-like protein is achieved by the transient induction of the inducible promoter in the host cell. In one embodiment of this type, the independent origin based cloning vector is a BBPAC and the recombination deficient host cell is an E. coli bacterium. In a preferred embodiment the RecA-like protein is recA.

BSPR:
The present invention also provides a conditional replication shuttle vector that encodes a RecA-like protein. In one such embodiment the RecA-like protein is controlled by an inducible promoter. In a preferred embodiment the conditional replication shuttle vector is a temperature sensitive shuttle vector (TSSV). The RecA-like protein of the TSSV can be controlled by either a constitutive promoter or by an inducible promoter. In one embodiment the TSSV contains a gene that can be counter-selected against. In a specific embodiment of this type the TSSV contains a gene that confers tetracycline resistance. In another embodiment the TSSV contains a RecA-like protein that is recA. In still another embodiment the TSSV contains both a gene that confers tetracycline resistance and a RecA-like protein that is recA. In a preferred embodiment the TSSV is pSV1.RecA having the ATCC no. 97968.

BSPR:
The present invention also provides an independent origin based cloning vector that contains a particular nucleotide sequence that has undergone homologous recombination with a conditional replication shuttle vector in a RecA-host cell, wherein the conditional replication shuttle vector encodes a RecA-like protein. In one such embodiment the particular nucleotide sequence is part of the gene that encodes the murine zinc finger gene, RU49 which is contained by the independent origin cloning vector. In one preferred embodiment the independent origin based cloning vector has undergone homologous recombination with a temperature sensitive shuttle vector in a RecA-host cell, wherein the temperature sensitive shuttle vector encodes a RecA-like protein. In another embodiment the independent origin based cloning vector is a BBPAC, and more preferably a BAC. In a specific embodiment of this type the independent origin based cloning vector has undergone homologous recombination with a temperature sensitive shuttle vector that is pSV1.RecA having the ATCC no. 97968.

BSPR:
In this aspect of the present invention, the independent origin based cloning vector contains a nucleic acid that has undergone homologous recombination with a conditional replication shuttle vector in a RecA.sup.- whole cell, in which the conditional replication shuttle vector includes a RecA like protein. In a preferred embodiment the independent origin based cloning vector is a BBPAC. In a more preferred embodiment, the BBPAC has undergone homologous recombination with a TSSV. In the most preferred embodiment, the BBPAC has undergone homologous recombination with the TSSV that is pSV1.RecA having the ATCC no. 97968.

BSPR:
One particular embodiment is a method of using the BBPAC to introduce the nucleic acid into an animal to make a transgenic animal comprising pronuclear injecting of the BBPAC (or a linearized nucleic acid insert derived from the BBPAC) into a fertilized zygote. In one embodiment the animal is a mammal. In a more preferred embodiment the mammal is a mouse. In a specific embodiment of this type the independent origin based cloning vector is a BBPAC and the fertilized zygote is a C57BL/6 mouse zygote. In a preferred embodiment of this type two picoliters (pl) of less than one .mu.g/ml BBPAC DNA is injected. In a more preferred embodiment 2 pl of 0.6 .mu.g/ml of DNA is injected.

BSPR:
The present invention also includes a method of using the BBPAC of the invention to perform gene targeting in a vertebrate cells comprising introducing the BBPAC into the vertebrate cell wherein the nucleic acid that has undergone homologous recombination with the conditional shuttle vector, undergoes homologous recombination with the endogenous chromosomal DNA of the vertebrate cell. In preferred embodiments of this type the vertebrate cell is a mammalian cell. In a more preferred embodiment of this type the mammalian cell is a human cell. In a related embodiment the vertebrate cell is a fertilized zygote and the nucleic acid contains a disrupted gene. In a preferred embodiment the conditional shuttle vector is a TSSV. In a more preferred embodiment the TSSV is pSV1.RecA having the ATCC no. 97968.

BSPR:
It is a further object of the present invention to provide a method of transiently expressing a RecA-like protein in a RecA.sup.- host cell to allow the specific modification of a gene of interest contained by an independent origin based cloning vector.

BSPR:
It is a further object of the present invention to provide a conditional replication shuttle vector which encodes a RecA-like protein, and which further contains a specific nucleic acid in a recombination cassette that selectively undergoes homologous recombination with an independent origin based cloning vector when both vectors are present in a recombination deficient host cell.

BSPR:
It is a further object of the present invention to provide a temperature dependent shuttle vector which encodes a RecA-like protein.

BSPR:
It is a further object of the present invention to provide a temperature dependent shuttle vector which encodes a RecA-like protein, which further contains a specific nucleic acid in a recombination cassette that can selectively undergo homologous recombination with a gene of interest contained by an independent origin based cloning vector, when both vectors are placed in a recombination deficient host cell.

DRPR:
FIGS. 2A and 2B show a schematic representation of targeted modifications of the BAC 169, which contains the murine zinc finger gene, RU49. BAC169 containing RU49 was obtained from screening of the mouse 129 strain BAC genomic DNA library (Research Genetics). FIG. 2A depicts a restriction map of the BAC169. The position of several exons are shown. The region of homology A1 (1 kb PCR fragment) and homology B1 (1.6 kb Xba-Hind fragment) are indicated. Abbreviations: XhoI (Xh), EcoRI (R), HindIII (H), XbaI (X), NotI (Not) and PmeI (Pme). FIG. 2B depicts a map of the modified BAC169 with IRES LacZ PolyA insertion (BAC169. ILPA). An extra PmeI site is inserted with the marker gene (asterisk). The size of the two Pme-Not fragments and the PmeI fragment are indicated. Since the marker gene (4 kb) is less than the deleted genomic region (7 kb), the total size of the modified BAC (128 kb) is smaller than the original BAC (131 kb).

DRPR:
FIGS. 4A-C show pulsed field gel electrophoresis analyses of modified 169 with the

ILPA insertion. DNA for two independent clones of BAC169. ILPA (L1 and L2) and BAC169 were prepared by alkaline lysis, and then digested with NotI, PmeI and XhoI (in a standard buffer supplemented with 2.5 mM spermidine). The digested DNA were separated by pulsed field gel electrophoresis (Bio-Rad's CHEF-DRII, 5 to 15s, 15 hours at 14.degree. C.) and blotted on to nitrocellulose filter (Stratagene). The same filter was probed separately with three probes. L1 and L2 are lacZ1 and LacZ2 which are independent clones which correspond to clones 1 and 2 respectively in FIGS. 3C and 3D.

DRPR:
FIG. 5A depicts purified linearized BAC L1 128 kb Not I insert for pronuclear injection. The pulsed field gel is probed with pgkpolyA probe. The numbers represent different fractions. The smear below the intact fragment represent degradation and undigested DNA.

DRPR:
FIG. 5B shows Southern blot analyses of the founder transgenic mice with the lacZ probe. The tail DNA were digested with Bam HI and Southern blot analysis was performed. The negative control consisted of littermates of Y3, Y7 and Y9 mice. The positive control was a conventional transgenic mouse with the lacZ transgene.

DRPR:
FIGS. 5C and 5D show the results of using PCR to determine the presence of BAC ends in the transgenic mice. The DNA at each end corresponding to the vector sequence is amplified and probed with a third oligonucleotide in the middle of the fragment. The appropriate size fragment is indicated. The negative controls are littermates. The positive control was BAC169 DNA.

DRPR:
FIG. 5E shows the germline transmission of the lacZ transgene in the Y7 mouse line. Tail DNA from two litters having eight mice each were prepared and digested with BamHI. Southern blot analysis was performed with the lacZ probe.

DEPR:
Targeted independent origin based cloning vector modification can be used for functional studies in diverse biological systems. The ability to efficiently modify a independent origin based cloning vector and generate an IOBCV-transgenic animal has important applications for functional analyses of genes in vivo. First, modified independent origin based cloning vectors can be used to study regulation of genes or gene complexes in transgenic animals such as mice. Since modified independent origin based cloning vectors can be used to study gene function in vivo, a deletion, substitution and point mutation within a given gene can be made in a independent origin based cloning vector, and the independent origin based cloning vector containing the modified gene can be reintroduced in vivo in its endogenous expression pattern. Furthermore, targeted independent origin based cloning vector modification can be used to create targeted expression of a selected gene, in the expression pattern of another gene, without prior knowledge of all of the regulatory elements of the selected gene. An important application of this type is targeted expression of the cre recominase for tissue/cell type specific gene targeting [Kuhn et al., Science 269:1427 (1995); Tsien et al., Cell 87:1317 (1996)]. Finally, modified independent origin based cloning vectors can be used to generate large DNA constructs particularly for gene targeting in ES cells and in vivo.

DEPR:
In one specific embodiment of the present invention the independent origin based cloning vector is a Bacterial Artificial Chromosome (BAC) modified in a host E. coli cell. A targeted BAC modification system has several advantages over a conventional yeast based modification system. First, a modified BAC automatically returns to the recombination deficient state after modification, ensuring stable maintenance of the modified BAC in the host strain. Second, BAC DNA can be very easily purified in relatively large quantities and high quality, allowing for use in biological experimentation including pronuclear injection. Third, since it is much easier to construct a BAC library than a YAC library, there are many more BAC libraries available from different species of animal, plants and microbes [Woo et al., Nucleic Acids Res., 22:4922 (1994); Wang et al., Genomics 24:527 (1994); Wooster et al., Nature 378:789 (1995)]. Most BACs also include all the necessary regulatory elements (i.e. LCRs and enhancers) to obtain dose dependent and integration site independent transgene expression [Dillon et al. Trends Genet. 9:134 (1993); Wilson et al., Annu. Rev. Cell. Biol. 6:679 (1990); Bradley et al., Nature Genet. 14:121 (1997)]. Targeted BAC modification can be applied successively to dissect these elements. In addition, such a modified BAC may be used to generate a transgenic animal. The BAC (or PAC) stably integrates into the animal cell genome. The transgenic animal can be used for functional studies, or for generating a desired gene product, such as producing a

human protein in the milk of a transgenic mammal [Drohan et al. U.S. Pat. No. 5,589,604, Issued Dec. 31, 1996]. Alternatively such modified BACs or PACs may be used for delivering a specific gene in gene therapy. In the Example below, a modified BAC has been successfully inserted into a murine subject animal, and in vivo heterologous gene expression has been demonstrated.

DEPR:
The methodology of the present invention is very general. Whereas the targeted independent origin based cloning vector modification is demonstrated on BACs, the system is readily applicable to BBPACs in general including PACs, P1 and other vectors propagated in the recombination deficient E. coli. In addition, the BAC modification exemplified herein, is also apropo to Mammalian Artificial Chromosomes. For example, Harrington et al. [Nature Genetics, 15:345-355 (1997)] have used BAC derived DNA as a component of their Human Artificial Chromosome. Therefore, the use of such human artificial chromosomes can include the BAC modification taught by the present invention.

DEPR:
In accordance with the present invention there may be employed conventional molecular biology, microbiology, and recombinant DNA techniques within the skill of the art. Such techniques are explained fully in the literature. See, e.g., Sambrook, Fritsch & Maniatis, Molecular Cloning. A Laboratory Manual, Second Edition (1989) Cold Spring Harbor aboratory Press, Cold Spring Harbor, N.Y. (herein "Sambrook et al., 1989"); DNA Cloning: A Practical Approach, Volumes I and II (D. N. Glover ed. 1985); Oligonucleotide Synthesis (M. J. Gait ed. 1984); Nucleic Acid Hybridization [B. D. Hames & S. J. Higgins eds. (1985)]; Transcription And Translation [B. D. Hames & S. J. Higgins, eds. (1984)]; Animal Cell Culture [R. I. Freshney, ed. (1986)]; Immobilized Cells And Enzymes [IRL Press, (1986)]; B. Perbal, A Practical Guide To Molecular Cloning (1984); F. M. Ausubel et al. (eds.), Current Protocols in Molecular Biology, John Wiley & Sons, Inc. (1994).

DEPR:
A "vector" is a replicon, such as plasmid, phage or cosmid, to which another DNA segment may be attached so as to bring about the replication of the attached segment. A "replicon" is any genetic element (e.g., plasmid, chromosome, virus) that functions as an autonomous unit of DNA replication in viva, i.e., capable of replication under its own control.

DEPR:
As used herein, a "Bacterial or Bacteriophage-Derived Artificial Chromosome" or "BBPAC" denotes a vector that is derived from a bacterium or bacteriophage such as a Bacterial Artificial Chromosome (BAC) which is an E. coli F element based cloning system, a P1-Derived Artificial Chromosome (PAC) or a lambda-based cosmid. In one embodiment, the BBPAC encodes up to 500 kilobases of genomic sequences. In a preferred embodiment, the BBPAC encodes between 120 to 180 kilobases of genomic sequences. In one particular embodiment the BBPAC encodes 130 kilobases of genomic sequences. A BBPAC used for gene targeting can be referred to as a "BBPAC targeting construct" and contains a nucleic acid insert comprising the gene targeting construct.

DEPR:
A "gene targeting construct" as used herein is used interchangeably with "targeting construct" and is a nucleic acid that when introduced into a cell undergoes homologous recombination with the endogenous chromosomal DNA of the cell. The nucleic acid is introduced into the cell to induce a modification of a particular gene contained on the endogenous chromosomal DNA, including in particular cases, to disrupt that gene to create a knockout animal.

DEPR:
As used herein a recombinant deficient host cell is "RecA.sup.- " when the host cell is unable to express a RecA-like protein, including recA itself, which can support homologous recombination. In the simplest case, the gene encoding the RecA-like protein has been deleted in a RecA.sup.- host cell. Alternatively the RecA-host cell contains a mutation in the recA gene that impairs its function.

DEPR:
A "RecA-like protein" is defined herein to have the meaning generally accepted in the art except as used herein the recA protein itself is included as being a specific RecA-like protein. RecA-like proteins are proteins involved in homologous recombination and are homologs to recA [Clark et al., Critical Reviews in Microbiology 20:125-142 (1994)]. The recA protein is the central enzyme in prokaryotic homologous recombination. It catalyzes pairing and strand exchange between homologous DNA molecules, and functions in both DNA repair and genetic

recombination [McKee et al., Chromosoma 7:479-488 (1996)]. A number of RecA-like proteins have been found in eukaryotic organisms and yeast [Reiss et al., Proc. Natl. Acad. Sci. 93:3094-3098 (1996)]. Two RecA-like proteins in yeast are Rad51 and Dmc1 [McKee et al. (1996) supra]. Rad51 is a highly conserved RecA-like protein in eukaryotes [Peakman et al., Proc. Natl. Acad. Sci. 93:10222-10227 (1996)].

DEPR:
A "cassette" refers to a segment of DNA that can be inserted into a vector at specific restriction sites. The segment of DNA encodes a polypeptide of interest, and the cassette and restriction sites are designed to ensure insertion of the cassette in the proper reading frame for transcription and translation. The present invention provides a recombination cassette that includes two homology fragments interrupted by an insertion, deletion or mutation sequence.

DEPR:
"Heterologous" DNA refers to DNA not naturally located in the cell, or in a chromosomal site of the cell. Preferably, the heterologous DNA includes a gene foreign to the cell.

DEPR:
A "nucleic acid molecule" refers to the phosphate ester polymeric form of ribonucleosides (adenosine, guanosine, uridine or cytidine; "RNA molecules") or deoxyribonucleosides (deoxyadenosine, deoxyguanosine, deoxythymidine, or deoxycytidine; "DNA molecules"), or any phosphoester analogues thereof, such as phosphorothioates and thioesters, in either single stranded form, or a double-stranded helix. Double stranded DNA-DNA, DNA-RNA and RNA-RNA helices are possible. The term nucleic acid molecule, and in particular DNA or RNA molecule, refers only to the primary and secondary structure of the molecule, and does not limit it to any particular tertiary forms. Thus, this term includes double-stranded DNA found, inter alia, in linear or circular DNA molecules (e.g., restriction fragments), plasmids, and chromosomes. In discussing the structure of particular double-stranded DNA molecules, sequences may be described herein according to the normal convention of giving only the sequence in the 5' to 3' direction along the nontranscribed strand of DNA (i.e., the strand having a sequence homologous to the mRNA). A "recombinant DNA molecule" is a DNA molecule that has undergone a molecular biological manipulation.

DEPR:
"Homologous recombination" refers to the insertion of a modified or foreign DNA sequence contained by a first vector into another DNA sequence contained in second vector, or a chromosome of a cell. The first vector targets a specific chromosomal site for homologous recombination. For specific homologous recombination, the first vector will contain sufficiently long regions of homology to sequences of the second vector or chromosome to allow complementary binding and incorporation of DNA from the first vector into the DNA of the second vector, or the chromosome. Longer regions of homology, and greater degrees of sequence similarity, may increase the efficiency of homologous recombination.

DEPR:
A DNA "coding sequence" is a double-stranded DNA sequence which is transcribed and translated into a polypeptide in a cell in vitro or in vivo when placed under the control of appropriate regulatory sequences. The boundaries of the coding sequence are determined by a start codon at the 5' (amino) terminus and a translation stop codon at the 3' (carboxyl) terminus. A coding sequence can include, but is not limited to, prokaryotic sequences, cDNA from eukaryotic mRNA, genomic DNA sequences from eukaryotic (e.g., mammalian) DNA, and even synthetic DNA sequences. If the coding sequence is intended or expression in a eukaryotic cell, a polyadenylation signal and transcription termination sequence will usually be located 3' to the coding sequence.

DEPR:
Transcriptional and translational control sequences are DNA regulatory sequences, such as promoters, enhancers, terminators, and the like, that provide for the expression of a coding sequence in a host cell. In eukaryotic cells, polyadenylation signals are control sequences.

DEPR:
A "promoter sequence" is a DNA regulatory region capable of binding RNA polymerase in a cell and initiating transcription of a downstream (3' direction) coding sequence. For purposes of defining the present invention, the promoter sequence is bounded at its 3' terminus by the transcription initiation site and extends upstream (5' direction) to include the minimum number of bases or elements necessary to initiate transcription at levels detectable above background. Within the promoter sequence

will be found a transcription initiation site (conveniently defined for example, by mapping with nuclease S1), as well as protein binding domains (consensus sequences) responsible for the binding of RNA polymerase.

DEPR:
A particular nucleotide sequence comprising a gene of interest, whether genomic DNA or cDNA, can be isolated from any source, particularly from a human cDNA or genomic library. In view and in conjunction with the present teachings, methods well known in the art, as described above can be used for obtaining such genes from any source (see, e.g., Sambrook et al., 1989, supra).

DEPR:
Accordingly, any animal cell potentially can serve as the nucleic acid source for the molecular cloning of any selected gene. The DNA may be obtained by standard procedures known in the art from cloned DNA (e.g., a DNA "library"), and preferably is obtained from a cDNA library prepared from tissues with high level expression of the protein by chemical synthesis, by cDNA cloning, or by the cloning of genomic DNA, or fragments thereof, purified from the desired cell (See, for example, Sambrook et al., 1989, supra; Glover, D. M. (ed.), 1985, DNA Cloning: A Practical Approach, MRL Press, Ltd., Oxford, U.K. Vol. 1, 11). Clones derived from genomic DNA may contain regulatory and intron DNA regions in addition to coding regions; clones derived from cDNA will not contain intron sequences.

DEPR:
In one aspect of the present invention, the recombination deficient host cell cannot independently support homologous recombination because the host cell is RecA.sup.-. However, as any person skilled in the art would readily understand, alternative causes for recombination deficiency may be rectified by methods that are analogous to those taught by the present invention and/or readily apparent in view of such teachings. For example recombination deficiency may be due to a deficiency of an alternative recombination protein such as another Rec protein including recB, recC, recD, and recE [Clark et al., Critical Reviews in Microbiol. 20:125-142 (1994)] which may be manipulated in a manner that is analogous to that taught herein for RecA-like proteins.

DEPR:
In the case of a RecA- host cell, inducing the host cell to transiently support homologous recombination comprises inducing the transient expression of a RecA-like protein in the host cell. Such induction may be performed by expressing a RecA-like protein contained by the recombination deficient host that is under the control of an inducible promoter.

DEPR:
In a preferred aspect of the invention inducing the transient expression of the RecA-like protein is performed with a conditional replication shuttle vector that encodes the RecA-like protein. Conditional replication shuttle vectors can also include pBR322 in a polyA temperature-sensitive bacterial strain. Preferably the conditional replication shuttle vector is a temperature sensitive shuttle vector (TSSV) that replicates at a permissive temperature, but does not replicate at a non-permissive temperature.

DEPR:
Inducing the transient expression of the RecA-like protein consists of transforming the host cell with the TSSV at a permissive temperature, and growing the host cell at a non-permissive temperature. The TSSV encodes a RecA-like protein that is expressed in the host cell and supports the homologous recombination between a specific nucleic acid contained in a recombination cassette and the particular nucleotide sequence contained in the host cell. The TSSV encoding the RecA-like protein is diluted out when the host cell is grown at the non-permissive temperature.

DEPR:
In a more intricate version of the present invention, the particular nucleotide sequence which has been selected to undergo homologous recombination is contained by an independent origin based cloning vector (IOBCV) that is comprised by the host cell, and neither the independent origin based cloning vector alone, nor the independent origin based cloning vector in combination with the host cell, can independently support homologous recombination. In a particular embodiment of this type both the independent origin based cloning vector and the host cell are RecA.sup.-, and inducing the host cell to transiently support homologous recombination comprises inducing the transient expression of the RecA-like protein to support homologous recombination in the host cell. The independent origin based cloning vector can be a BBPAC, such as the BAC exemplified below and the host cell can be a host bacterium, such as E. coli.

DEPR:
The independent origin based cloning vectors for use in the methods of the present invention can be obtained from a number of sources. For example, E. coli-based artificial chromosomes for human libraries have been described [Shizuya et al., Proc. Natl. Acad. Sci. 89:8794-8797 (1992); Ioannou et al., In Current Protocols in Human Genetics (ed. Dracopoli et al.) 5.15.1-5.15.24 John Wiley & Sons, New York (1996); Kim et al., Genomics 34:213-218 (1996)]. Libraries of PACs and BACs have been constructed [reviewed in Monaco et al., Trends Biotechol., 12:280-286 (1994)], that are readily isolated from the host genomic background for example by classical alkaline lysis plasmid preparation protocols [Birnboim et al., Nucleic Acids Res. 7:1513-1523 (1979)], or alternatively, with the use of a nucleobond kit, a boiling Prep, or by cessium gradient (Maniatis, supra). BAC, PAC, and P1 libraries are also available for a variety of species (e.g. Research Genetics, Inc., Genome Research, Inc., Texas A&M has a BAC center to make a BAC library for livestock and important crops). Also BACs can be used as a component of mammalian artificial chromosomes.

DEPR:
An independent origin based cloning vector that is a BAC can be isolated using a cDNA or genomic DNA probe to screen a BAC genomic DNA library, for example. The use of a mouse genomic BAC library from Research Genetics is exemplified below. A positive BAC can generally be obtained in a few days. To insert a gene of interest into a selected locus in the BAC, the region of insertion can be mapped for restriction enzyme sites. Whereas subcloning is necessary for detailed mapping, it is generally unnecessary since rough mapping is usually sufficient. As is readily apparent, other independent origin based cloning vector genomic libraries can be screened and the isolated independent origin based cloning vectors manipulated in an analogous fashion.

DEPR:
The conditional replication shuttle vectors of the present invention are constructed so as to contain a recombination cassette that can selectively integrate into the nucleotide sequence of the gene of interest encoded by the independent origin based cloning vector. Such conditional replication shuttle vectors can be constructed by inserting a PCR amplified RecA-like gene into an appropriate conditional replication shuttle vector which either contains a specific drug resistant gene or can be subsequently modified to contain one. In a preferred embodiment the drug resistant gene can also be counter-selected against, such as with, tetracycline and fusaric acid. Alternatively, in addition to the drug resistant gene the conditional shuttle vector can also contain a counter-selection gene such as a gene that confers sensitivity to galactose, for example.

DEPR:
In the Example below, the E. coli K12 recA gene (1.3 kb) is inserted into the BamHI site of a pMBO96 vector. In this case the vector already carried a gene that bestows tetracycline resistance, and in addition contains a pSC101 temperature sensitive origin of replication, which allows the plasmid to replicate at 30 degrees but not at 43 degrees celsius.

DEPR:
The RecA-like protein of a conditional replication shuttle vector can be controlled by either an inducible promoter or a constitutive promoter. In one particular embodiment the transient expression of the RecA-like protein is achieved by the transient induction of the inducible promoter in a host cell. In another embodiment, the constitutive promoter is the endogenous E. coli recA promoter.

DEPR:
In certain cases a building vector is used to construct the recombination cassette. Two small genomic fragments, each containing about 500 basepairs (400 basepairs to 600 basepairs is sufficient) of the gene of interest are cloned into the building vector (e.g., pBV1) in appropriate order and orientation to generate the flanking regions of the recombination cassette. DNA containing a promoter sequence 5' to the specific nucleic acid, which in turn is 5' to a polyadenine addition signal sequence, is inserted between the two genomic fragments in the proper orientation. The recombination cassette is then transferred into the conditional replication shuttle vector (e.g., pSV1.RecA). The recombination cassette, the RecA-like protein gene, and the drug resistant gene are linked together on the conditional replication shuttle vector such that when the specific nucleic acid integrates into the particular nucleotide sequence, the RecA-like protein gene and the drug resistant gene remain linked together, and neither the RecA-like protein gene nor the drug resistant gene remain linked to the integrated specific nucleic acid. In a preferred embodiment the conditional replication shuttle vector is a TSSV and the TSSV is pSV1.RecA having the ATCC no. 97968, which has been deposited with the American Type Culture Collection, 10801 University Boulevard, Manassas, Va., 20110-2209, on Mar. 25, 1997 under the

Budapest Treaty.

DEPR:
According to the methods of the present invention the conditional replication shuttle vector is transformed into a RecA.sup.- host cell containing the independent origin based cloning vector. The independent origin based cloning vector can also contain a gene which bestows resistance to a host cell against a corresponding toxic agent/drug such as an antibiotic or in a specific embodiment, chloramphenicol. The cells are grown under the conditions in which the conditional replication shuttle vector can replicate (e.g., when the conditional replication shuttle vector is a TSSV which replicates at 30.degree. but not at 43.degree., the host cell is grown at 30.degree. C.) and the transformants can be selected via the specific drug resistant gene (or first drug resistant gene) carried by conditional replication shuttle vector, and the second drug resistant gene carried by the independent origin based cloning vector. Since the conditional replication shuttle vector also carries the RecA-like protein gene, homologous recombination can occur between the conditional replication shuttle vector and the independent origin based cloning vector to form co-integrates through the sequence homology at either the 5' or the 3' flanking regions of the recombination cassette. The co-integrates then can be selected by growing the cells on plates containing the first and second drugs at non-permissive conditions (e.g. for the TSSV above, at 43.degree. C.) so that the non-integrated, free conditional replication shuttle vectors are lost. This results in the selection for host cells carrying the integrated conditional replication shuttle vectors, (which co-integrate either into the independent origin based cloning vector or into the host chromosome). Correct independent origin based cloning vector co-integrates can be identified by PCR or more preferably with Southern blot analyses.

DEPR:
The co-integrates can then be re-streaked onto plates containing the second drug, (i.e., the drug which the gene initially carried by the independent origin based cloning vector protects against) and grown under non-permissive conditions overnight. A fraction of the co-integrates undergo a second recombination event (defined as resolution), through sequence homology at either the 5' or the 3' flanking regions of the recombination cassette. The resolved independent origin based cloning vector automatically loses both the first drug resistant gene (i.e., the specific drug resistant gene contained by the conditional replication shuttle vector) and the RecA-like protein gene due to the linkage arrangement of the RecA-like protein gene, the drug resistant gene and the specific nucleic acid on the conditional replication shuttle vector, described above. In addition, the excised conditional replication shuttle vector cannot replicate under the non-permissive conditions and is therefore diluted out.

DEPR:
The resolved independent origin based cloning vectors can be further selected for by growing the host cells (e.g., at 37.degree. C.) on plates containing the second drug and an agent that counterselects against cells containing the gene resistant to the first drug (e.g., a gene conferring tetracycline resistance may be counter-selected against with fusaric acid). The resolved independent origin based cloning vector will be either the original independent origin based cloning vector or the precisely modified independent origin based cloning vector. One method to identify the correctly resolved BAC is to choose 5-10 colonies and prepare a miniprep DNA. The DNA can then be analyzed using Southern blots to detect the correct targeting events. Alternatively, the desired clones can be identified by colony hybridization using a labeled probe for the specific nucleic acid contained by the recombination cassette. Such probes are well known in the art, and include labeled nucleotides probes that hybridize to the nucleic acid sequence. Alternatively, a marker nucleic acid can be included in the recombination cassette and constructed so as to remain with the specific nucleic acid upon integration into the independent origin based cloning vector.

DEPR:
In an alternative method, the modified independent origin based cloning vector and the unmodified independent origin based cloning vector can be assayed with both a probe specific for any region of the DNA contained by the recombination cassette predicted to be inserted into the independent origin based cloning vector (e.g., the promoter sequence, the specific nucleic acid, and a polyadenine addition signal sequence) and a probe specific for a region outside of the modification region (e.g., near the promoter region but outside of the modification region).

DEPR:
A modified independent origin based cloning vector of the present invention can be purified by gel filtration, e.g. a column filled with SEPHAROSE CL-4B yielded intact linear BAC DNA. The column can be pre-equilibrated in an appropriate buffer, as

described in the Example below. The purified DNA can be directly visualized with ultraviolet light after ethidium bromide staining, for example. Columns such as the SEPHAROSE CL-4B column also can efficiently separate degraded DNA from the pure linear DNA.

DEPR:
The present invention also provides methods of using the modified independent origin based cloning vectors of the present invention. Such modified independent origin based cloning vectors contain a nucleic acid that can be inserted into an animal to make a transgenic animal. The modified independent origin based cloning vectors of the present invention can be introduced into the desired host cells by methods known in the art, e.g., transfection, electroporation, microinjection, transduction, cell fusion, DEAE dextran, calcium phosphate precipitation, lipofection (lysosome fusion), use of a gene gun, or a DNA vector transporter (see, e.g., Wu et al., 1992, J. Biol. Chem. 267:963-967; Wu and Wu, 1988, J. Biol. Chem. 263:14621-14624; Hartmut et al., Canadian Patent Application No. 2,012,311, filed Mar. 15, 1990).

DEPR:
One particular method comprises the pronuclear injection of the modified independent origin based cloning vector into a fertilized animal zygote. Such a method is exemplified below with the modified independent origin based cloning vector being a BAC which has been linearized, and the animal zygote being a mouse zygote. 2 pl of 0.6 .mu.g/ml of BAC DNA was injected.

DEPR:
The presence of both ends of the modified independent origin based cloning vector can be assayed for in the transgenic animal to determine if the intact nucleic acid insert of the IOBCV has been integrated into the genome. Since both ends of the nucleic acid insert contain some vector sequence, PCR primers specific to the vector sequence can be generated and used to amplify the transgenic DNA. The amplified products can then be probed with a third labeled oligonucleotide probe within the amplified region.

DEPR:
A preferred version of the BBPAC gene targeting methodology of the present invention also includes negative selection. The conventional negative selection cassettes, such as the use of the herpes thymidine kinase cassette or the diphtheria toxin gene cassette, may not always work with BBPAC constructs since BBPAC DNA tends to exist in transfected mammalian cells as episomal DNA for a long period of time [Baker et al., NAR 25:1950-1956]. In one example, the EGFP1 cassette can be used as a negative screening cassette. In this case, in the second step of modification to generate the short arm, the CMV promoter driven green fluorescent protein (EGFP-1) and the polyA signal can be introduced. Unlike other negative selection cassettes, GFP is not toxic to the cells but serves as a fluorescent marker protein. When gene targeting occurs, the EGFP-1 cassette will be lost and the cell will not exhibit a green fluorescence under UV light. On the other hand, when the BBPAC integrates non-homologously, the EGFP-1 cassette also integrates, and the cells will therefore exhibit the green fluorescence under UV. For the definitive Southern blot analyses only those neo resistant cell lines which do not exhibit a green fluorescence under UV light are chosen.

DEPR:
The method enables transfection efficiency in mammalian cells with linear BBPAC DNA to be similar to the transfection efficiency of a conventional DNA construct. On the other hand, the BBPAC targeting construct can potentially provide 10-100 fold higher targeting frequency than the conventional targeting construct, thereby making gene targeting in mouse ES cells easier and cheaper, since only a few dozen colonies need to be isolated and screened to obtain the targeted clones.

DEPR:
The present invention further provides a method of performing gene targeting in fertilized vertebrate zygotes by the injection of a BBPAC targeting construct, or preferably the linearized intact BBPAC nucleic acid insert containing the targeting construct to generate a transgenic knock-out animal (TKO). A large targeting construct (>100 kb) can provide a very high targeting rate (predicted by mathematical modeling described above) and gene targeting can be directly performed with a fertilized vertebrate zygote via pronuclear injection of the modified BBPAC targeting construct. TKO methodology has previously been attempted by Brinster et al. [PNAS, 86:7087-91 (1989)] with a small DNA construct (2.6-8.9 kb) but those workers only obtained a relatively low targeting rate (0.2%). The large homology DNA in the BBPAC (>100 kb) of the present invention increases the targeting rate to a favorable range of 2% to 10%.

DEPR:
The modified BBPAC TKO construct can be prepared in milligram quantities and linearized as described above. The linearized DNA then is introduced into the fertilized zygote by a standard protocol, e.g., pronuclear injection (Hogan et al., (1986) supra). The transgenic animal is then identified by standard Southern blots. The gene targeting event can be further identified by digesting DNA of the transgenic animal with appropriate enzymes, such as enzyme X, (FIG. 7) and probed with the flanking KO probe (FIG. 7). Mice with the targeting event will have an additional band of the appropriate size. Such gene targeting events can further be confirmed by expression of the GFP or LacZ marker gene in the expression pattern of the targeted endogenous gene, since the construct is designed to trap the endogenous promoter.

DEPR:
In still another aspect of the present invention, methods of performing gene targeting in somatic cells using BBPAC targeting constructs are provided. Since gene targeting in somatic cells is also dependent on the length of homology, using large DNA targeting construct also improves the targeting rate in somatic cells. The experimental design in this case is similar to that with the ES cells described above. Somatic cell gene targeting is useful in gene therapy, for example, in a targeted insertion of a functional gene in a hereditary disease of the hematopoietic system. Such methods are also useful to generate targeted cell lines for experimental purposes.

DEPR:
Conditional replication shuttle vectors that encode a RecA-like protein are also provided by the present invention. The RecA-like protein can be controlled by either an inducible promoter or a constitutive promoter. The conditional replication shuttle vector is preferably a temperature sensitive shuttle vector (TSSV). In one such embodiment the TSSV contains both a gene that confers tetracycline resistance and a RecA-like protein that is recA. In a preferred embodiment, the TSSV is pSV1.RecA having the ATCC no. 97968.

DEPR:
Independent origin based cloning vectors that contain a gene of interest that has been modified by the methods of the present invention are also included in the present invention. More particularly such independent origin based cloning vectors have undergone homologous recombination with a conditional replication shuttle vector in a RecA.sup.- host cell, wherein the conditional replication shuttle vector encodes a RecA-like protein. In a preferred embodiment the independent origin based cloning vector has undergone homologous recombination in a RecA.sup.- host cell with a temperature sensitive shuttle vector encoding a RecA-like protein. In a preferred embodiment the modified independent origin based cloning vector is a BAC that has undergone homologous recombination with the temperature sensitive shuttle vector pSV1.RecA having the ATCC no. 97968.

DEPR:
Bacterial based artificial chromosomes, such as Bacterial artificial chromosomes (BACs) and P-1derived artificial chromosomes (PACs), are circular bacterial plasmids that may propogate as large as 300 kb of exogenous genomic DNA (Shizuya et al, PNAS, 89, 8794-97, 1992; loannou et al, Nature Genet., 6, 84-90, 1994). For the majority of BAC and PAC libraries, the average size of the insert is 130-150 kb. There are several advantages of using bacterial based artificial chromosomes for genomic and functional studies, compared to the yeast based system (i.e. YACs): First, BAC and PAC libraries are much easier to construct due to higher cloning efficiency. Second, BACs and PACs are propagated in recombination deficient E. coli host cells, so they have high stability and minimal chimerism. No rearrangements have been observed in BACs or PACs after 100 generations of growth. Third, isolation of BAC and PAC DNA is very easy since they exist as supercoiled circular plasmids that are resistant to shearing. Conventional bacterial plasmid DNA isolation methods can be applied to obtain milligrams of intact BAC or PAC DNA. Finally, direct DNA sequencing can be applied to BAC or PAC DNA, which is not possible for YAC DNA. These advantages have made BACs and PACs important tools for genome studies in many species.

DEPR:
Although BBPACs are useful for physical mapping in genome studies, no simple method is available to modify BBPACs, as is available for the YACs. A simple homologous recombination based BBPAC modification method is disclosed, termed targeted BBPAC modification (See FIG. 7 for a schematic representation of the method). This method allows precise modification, such as marker insertion, deletion, point mutation, at any chosen site within a given BBPAC. This method involves several steps: isolation of BBPACs using cDNA or genomic DNA probes, simple mapping and partial sequencing of the BBPACs, cloning of the shuttle vector, targeted modifications, pulsed field gel analyses of the modified BBPACs, and finally preparation of linearized BBPAC DNA for

functional studies, such as pronuclear injection to produce BBPAC transgenic mice. Since the method is simple and reliable, it is reasonable to expect that the entire procedure, from the step of screening for a BBPAC with a cDNA or genomic DNA probe to the step of modified BBPACs ready for functional studies, can be completed within 6-8 weeks.

DEPR:
A BAC clone is isolated with either a unique cDNA or genomic DNA probe. BAC libraries for various species, (in the form of high density BAC colony DNA membrane) are available from Research Genetics, Inc. and Genome Research, Inc. The mouse 129 genomic BAC library from Research Genetics has proved to be a good source for genomic DNAs. To avoid damage to the membrane, the probe is first tested on a mouse genomic Southern blot to ensure that the probe does not contain any repetitive elements. The library is screened according to manufacture's direction. The positive clones can be obtained from the company within a few days.

DEPR:
(II) Preparation of midiprep BAC DNA by alkaline lysis method (1 day):

DEPR:
(III) BAC maxiprep DNA preparation:

DEPR:
Two methods were used to prepare large quantities of RNA-free BAC maxiprep DNA. The first method is the standard cesium chloride banding method (see Maniatis, supra). This method was used routinely to obtain >500 ug BAC DNA from 1 liter bacteria culture. The second method, uses a commercially available column, the Nucleobond AX-500 (made by The Nest Group, Southborough, Mass.). The maxiprep DNA are also stored in 4.degree. C. for long-term storage.

DEPR:
To determine the size of each BAC and to confirm that the BAC contains the gene of interest, a simple mapping of the BACs is done. The following enzymes are used to map each BAC: Not I (to release the BAC insert), Mlu I, NotI/Mlu I (double digest), PmeI, PmeI/NotI and XhoI. Digestion is done in a 40 ul total volume, which contains the following: 5 ul midiprep DNA, 4 ul digestion buffer, 4 ul 10.times.BSA (if necessary), 1 ul 100 mM spermidine (final concentration 2.5 mM), 2 ul enzyme (10-40 units), and ddH2O. Digestion is done at 37.degree. C. for >5 hrs.

DEPR:
The digested BACs are resolved on a pulsed field gel (Bio-Rad's CHEF-DRII). The gel is 1% agarose in 0.5.times.TBE. The gel is run in 0.5.times.TBE. The separation condition is the following: 6v/cm, 5s to 15s linear ramping for 15 hrs to 18 hrs at 14.degree. C. The New England Biolab's PFGE marker I or II as the high molecular weight marker and 1 kb DNA ladder (Life Technologies Inc.) as the low molecular weight marker are used.

DEPR:
The gel is then stained with ethidium bromide (1 to 5000, or 1 to 10,000 dilution of 10 mg/ml stock) for 30 min prior to taking the photograph. Then the gel is blotted onto the nitrocellulose membrane and hybridized to cDNA and genomic DNA probes according to standard protocols (Maniatis, supra). To ensure the entire cDNA is included in the BAC, probes/or oligonucleotides from both the 5' end and the 3' end of the gene are used to probe the blot separately. Those large BACs containing the entire gene are usually selected for BAC modification.

DEPR:
Each sequencing reaction will result in up to a 500 bp sequence. Sequence more than one BAC for a given primer to compare the sequences. The main purpose for sequencing is to design a 20 bp PCR primer, which is about 500 bp away from the sequencing oligo (which usually is the other PCR primer), to enable PCR amplification of this genomic fragment and to clone it into the building vector. Therefore, as long as a 20 bp sequence can be identified which is at the appropriate position, and which is the same in several independent sequencing reactions, the goal is achieved. The quality of the DNA sequence in between is not very critical.

DEPR:
A two vector system is designed to construct the shuttle vector for BAC modification (FIG. 1). The first vector is a pBS.KS based building vector, which is used to construct the recombination cassette containing homologous sequence A and homologous sequence B and the modification to be introduced between them. The recombination cassette was not constructed in the pSV1.RecA shuttle vector was for the following reasons: first, it is a low copy plasmid so that it is difficult to obtain high

quantity DNA; second, it is a large plasmid (11 kb), so it is relatively difficult to clone. The building vector contains the marker gene to be introduced into the BAC, cloning sites flanking it (usually EcoRI for cloning the homology A and XbaI for homology B, and rare restriction sites such as MluI, PmeI and Pac I for mapping of the modified BAC). There are two Sal I sites (or one Sal I, one XhoI) flanking the multiple cloning sites. They are used to release the recombination cassette and subclone it into the Sal I site of the pSV1.RecA vector, to complete the shuttle vector construction. One thing about designing the building vector is that there should not be any Not I sites within the recombination cassette, since NotI sites are used in the end to release the linear modified BAC for biological experiment (e.g., pronuclear injection). The map and utility of various building vectors and the shuttle vector are described below.

DEPR:
(II) Prepare intact linearized BAC DNA for pronuclear injection (1 day):

DEPR:
To modify BACs in E. coli, a temperature sensitive shuttle vector based system for homologous recombination was employed [O'Connor et al., Science 244:1307-1312 (1989); Hamilton et al., J. Bacteriol. 171:4617 (1989)]. This temperature sensitive plasmid will replicate in cells growing at the permissive temperature (30.degree. C.), but will be lost in cells growing at the restrictive temperature (42-44.degree. C.) because its origin of replication can not function at the restrictive temperature [Hashimoto-Gotoh et al., J. Bacteriol. 131:405-412 (1977)]. To overcome the recombination deficiency of the BAC host i.e., a RecA.sup.- host cell, the E. coli recA gene was introduced into the temperature sensitive shuttle vector. When transformed with the temperature sensitive shuttle vector (carrying a recombination cassette containing the recA gene) the host strain becomes conditionally competent to perform homologous recombination allowing in vivo modification of the resident BAC.

DEPR:
The general strategy for targeted BAC modification is shown in FIG. 1, which illustrates the steps involved in inserting a marker gene, e.g., IRES-lacZ-pGKpolyA (ILPA), into the BAC. First, two small genomic fragments, e.g., A and B, each containing greater than 500 basepairs of a gene of interest are cloned into the building vector (pBV1) in appropriate order and orientation to generate the recombination cassette. The recombination cassette is then transferred into the temperature sensitive shuttle vector (e.g., pSV1.RecA). The reason the recombination cassette is not built directly in the shuttle vector is due to the relative difficulty in manipulating its DNA, due to low copy number [Bochner et al., J. Bacteriol. 143:926 (1980); Maloy et al., Bacteriol. 145:1110 (1981)] and large vector size (11 kb).

DEPR:
This shuttle vector is then transformed into E. coli containing the BAC. The transformants can be selected by tetracycline resistance (carried by pSV1.RecA) and chloramphenicol resistance (carried by the BACs) at 30.degree. C. Since the shuttle vector also carries the recA gene, homologous recombination can occur between the shuttle vector and the BAC, through either homology at A or B to form co-integrates. The co-integrates are selected by growth on tetracycline and chloramphenicol plates at 43.degree. C. This temperature is non-permissive for shuttle vector replication, so that the non-integrated, free shuttle vectors are lost, resulting in the selection for bacteria carrying the integrated shuttle vectors, (either into the BACs or into the bacterial chromosomes). Correct BAC co-integrates can be identified by Southern blot analyses.

DEPR:
The co-integrates are then restreaked onto the chloramphenicol plates and grown at 43.degree. C. overnight. A fraction of the co-integrates will undergo a second recombination event (resolution), through either homology at A or B. The resolved BACs will automatically lose the tet and the recA genes, since the excised shuttle vector plasmids cannot replicate at the non-permissive temperature. The resolved BACs can be selected by growing on chloramphenicol and fusaric acid plates at 37.degree. C., as growth on fusaric acid plates selects for the loss of tetracycline resistance, i.e., counterselecting against BACs that are resistant to tetracycline. As illustrated in FIG. 1, depending on which pair of homologous fragments undergo the second recombination event, the resolved BAC can be either the original BAC or the precisely modified BAC. The desired clones can be identified by colony hybridization using a labeled probe for the inserted marker. One important aspect of the method is that the recA gene is only temporally introduced into the bacterial host. Once the modification is finished, the bacteria will automatically lose the recA gene, returning to the recombination deficient state suitable for stable maintenance of the modified BACs.

DEPR:
This strategy termed targeted modification of BACs, was tested by introducing the IRES-lacZ-polyA (ILPA) marker into the 131 kb murine BAC169 containing the RU49 locus (FIG. 2A). In this case, the marker gene to the first coding exon of the RU49 gene was targeted with homology fragments being 1 kb and 1.6 kb respectively (FIG. 2B). Placing the IRES sequence before the lacZ gene ensures the translation of the marker gene even when lacZ gene is placed after the translation start site [Pelletier et al., Nature 334:320 (1988)]. The pSV1.RecA temperature sensitive shuttle vector containing the recombination cassette was transformed into the DH10 E. coli strain containing the BAC169 and selected by growth at either 30.degree. C. or 43.degree. C. on plates containing chloramphenicol and tetracycline. In contrast to growth at 30.degree. C., which produced a thick lawn of transformed cells, growth at 43.degree. C. resulted in growth of individual colonies. Twenty of these were picked and tested by Southern blots for co-integration of the shuttle plasmid into BAC169. As shown in FIG. 3B, analysis of twenty clones using the B1 fragment of the RU49 homology cassette resulted in the identification of two clones containing the appropriate 4 and 8 kb EcoRI bands (10%), indicating that these clones carry co-integrates that have occurred through this region of homology.

DEPR:
The next step in our analysis was extensive mapping of the modified BACs to determine whether any unexpected deletions or insertions were generated during the modification procedure. FIG. 4 shows pulsed field gel mapping of the modified BAC L1 and L2 and the original BAC 169. The same filter was probed separately with the whole BAC169 probe, with a probe from the inserted marker gene (pgkpolyA) and a probe from the 5' non-modified region of the RU49 gene (A2). BAC169 probe (left panel) hybridizes with all the restriction fragments for each BAC. Thus, XhoI digestion reveals a finger print of the modified BACs showing that essentially all fragments are preserved. The only difference is that the fragment containing the ILPA insert is slightly smaller than the corresponding wild type fragment due to the replacement of the 7 kb RU49 fragment with the 4 kb marker gene (FIG. 2B). Digestion with NotI, which releases the entire BAC insert, also reveals a slightly smaller DNA insert in modified BACs for the same reason. Since the marker gene was engineered to carry an additional PmeI site (FIG. 2), digestion of the BAC L1 and L2 DNAs with this enzyme results in the generation of two fragments, in contrast to the single fragment seen in the original BAC69. The sizes of these fragments allow the determination that these BACs contain approximately 75 kb 5' to the PmeI site, and 53 kb 3' to it (FIG. 2). No apparent rearrangements have occurred during the modification procedure.

DEPR:
To confirm this conclusion, the modified BACs and BAC169 were probed with both a marker specific probe (pgkpolyA) and a probe near the promoter region and outside the modification region (A2). Consistently, both modified BACs contained a single band homologous to the marker gene probe which is not present in BAC169. When the A2 probe was used, a single band of expected size appeared in all three BACs. Additional fingerprinting of all eight modified BACs with HindIII, EcoRI and AvrII digests showed that no detectable rearrangements or deletion existed in these BACs. Thus, the temporary introduction of the recA gene into the BAC host strain does not introduce any rearrangements or deletions.

DEPR:
To demonstrate the possibility of using the modified BACs for in vivo studies for gene expression and gene function, transgenic mice carrying the modified BAC169 with the IRES-LacZ insertion were generated. To purify the 128 kb BAC insert for pronuclear injection, several established methods for purifying large YAC DNA were attempted, and resulted in considerable amount of DNA fragmentation. In contrast, when a simple gel filtration column filled with SEPHAROSE CL-4B was tried, very pure fractions of intact linear BAC DNA insert were obtained in an appropriate injection buffer, e.g., 100 mM NaCl, 10 mM Tris.HCl, pH 7.5 and 0.1 mM EDTA (FIG. 5A). Unlike YAC DNA purification which typically results in a low DNA yield, the purified fractions using the SEPHAROSE CL-4B column contained a large quantity of high concentration linear DNA (e.g., 0.5 mls of 3 .mu.g/ml DNA or more). The purified DNA could be directly visualized with ultraviolet light after ethidium bromide staining. The SEPHAROSE CL-4B column could also efficiently separate the degraded DNA (in this case in fractions 3-6) from the pure linear DNA (fractions 7-9) (FIG. 5A). Fraction 8 contained 3 .mu.g/ml DNA and was used directly for pronuclear injection.

DEPR:
Pronuclear injection into the fertilized C57BL/6 mouse zygote is performed according to a standard protocol [Hogan et al., in Manipulating the Mouse Embryo (Cold Spring Harbor Laboratory Press, New York, 1986)]. Two different concentrations of fraction 8 BAC DNA (obtained as described above) were used: 3 .mu.g/ml and 0.6 .mu.g/ml. No

newborns were obtained with the high concentration DNA, suggesting that the high concentrations may be toxic to the zygote. However, with the lower concentration of pure linear DNA, 15 newborn mice were obtained and two of them (13%), Y7 and Y9, contained the lacZ marker gene as demonstrated on a Southern blot (FIG. 5B). The intensity of the bands allows an estimate of 2-3 transgene copies for Y7 and one copy for Y9.

DEPR:
To determine if the intact BACs have been integrated into the genome, the presence of both ends of the BAC ends was assayed for in the transgenic mice. Since both BAC ends contain some vector sequence, PCR primers specific to the vector sequence were generated and used to amplify the transgenic DNA. The amplified products were then probed with a third labeled oligonucleotide probe within the amplified region. As shown in FIG. 5C and FIG. 5D: Y3, Y7 and Y9 have both ends present, while the negative controls do not. Since Y7 and Y9 also have the lacZ gene, they are likely to contain intact BAC transgenes. For Y3, whereas it has both ends it does not contain the lacZ gene. This may be due to either a rearrangement or fragmentation during the injection prior to integration.

DEPR:
Next the expression of lacZ gene in the cerebellum of the Y7 transgenic mice was determined by whole mount lacZ staining. RU49 is normally expressed in the granule cells of the cerebellum, the dentate gyrus and the olfactory bulb (including the subventricular zone, the rostral migratory stream, and the olfactory bulb proper) [Yang et al., Development, 122:555-566 (1996)]. In previous studies, RU49 promoter lacZ transgenic mice with 10 kb promoter had been generated. However, all of the transgenic lines showed strong positional effects: either they did not express in the brain at all, or they were ectopically expressed in the cortex, but not the cerebellum. One particular 10 kb-lacZ transgenic line did show restricted expression in the cerebellum, however, the expression was restricted to the caudal half of the cerebellum. With 128 kb of RU49 endogenous sequence surrounding the lacZ gene in the Y7 line, at postnatal day 6, the transgenic mice showed a lacZ expression pattern closely resembling the endogenous expression pattern (FIG. 6). In the cerebellum, the marker gene is expressed throughout the cerebellum (FIG. 6A) and no expression is seen in five control littermates (FIG. 6B). Further analysis showed that the transgene is expressed at high level in the EGL and lower level in the IGL. The lacZ marker gene is also expressed in the dentate gyrus and the rostral migratory stream and the olfactory bulb (FIGS. 6C and 6D). The pattern of the BAC transgene expression closely resembles the endogenous RU49 expression pattern in the brain. It is evident that the large genomic DNA in the BAC transgene can overcome the positional effects and confer the proper expression of RU49 in vivo, in contrast to our results using conventional transgenic constructs.

DEPR:
As taught herein, bacterial based artificial chromosomes (BACs and PACs) are ideal for constructing large DNA for gene targeting. As demonstrated herein with the targeted BAC modification method, BACs and PACs can be readily modified to introduce selection genes, marker genes, and deletions. Making a BBPAC gene targeting construct will take about the same time as making a conventional targeting construct (1-3 months). Moreover, BBPAC targeting construct DNA can be easily isolated in milligram quantity and high quality. This is advantageous over the YAC system, since it is difficult to purify large quantities of high quality YAC DNA.

DEPL:
4. Preparation of Large Quantity, High Quality Linearized BAC DNA for Pronuclear Injection

DEPC:
HOMOLOGOUS RECOMBINATION BASED MODIFICATION IN E. COLI AND GERMLINE TRANSMISSION IN TRANSGENIC MICE OF AN 131 KILOBASE BACTERIAL ARTIFICIAL CHROMOSOME

DEPV:
9). Wash the pellet with 70% ethanol. Dry by vacuum and resuspend the DNA in 200 ul TE. The BAC midiprep DNA may be stored in 4.degree. C. for months (Do not freeze the BAC DNA, since repetitive freezing and thawing will result in degradations).

DEPV:
1) If maxiprep DNA is used, go directly to step 2. If midiprep DNA is used, first add 100 ul ddH2O and 10 ul 10 mg/ml RNAse A to 100 ul midiprep BAC DNA, and incubate at 37.degree. C. for >1 hr. (This step is critical, incomplete RNAse treatment will result in poor precipitation and sequencing).

DEPV:

2) Add 132 ul PEG mix (2.5M NaCl and 20% PEG 8000) to the treated DNA. Put on ice for 5 min.

DEPV:
5) Run 2 ul on a agarose gel to estimate the final concentration. Usually use 6-8 ul (500 ng-1000 ng) DNA for automatic sequencing, also use 150 ng sequencing oligos.

DEPV:
This plasmid vector was modified from the pMBO96 vector originally constructed by O'Connor et al (Science, 1989, Vol 244, pp. 1307-1312). The pMBO96 vector was a gift from Dr. Michael O'Connor. The original vector carries tetracycline resistance, and contains a pSC101 temperature sensitive origin of replication, which allows the plasmid to replicate at 30.degree. C. but it will cease replication and is lost at 43.degree. C. The E. coli RecA gene was amplified by PCR and sub-cloned into the Bam HI site, to create the pSV1.RecA vector. The Sal I site is used to subclone the recombination cassette from the building vector.

DEPV:
2. Prepare pSV1.RecA and building vector midi-prep DNA by the alkaline lysis method (see above). For the pSV1.RecA vector, Qiagen columns can also be used to obtain high purity DNA, though yield is usually low. This is due to the low copy number of the pSV1 plasmid. For preparation of pSV1.RecA DNA, the culture should be grown at 30.degree. C. in LB+tetracycline (10 ug/ml). The final midi-prep DNA is usually dissolved in 200 ul TE or ddH.sub.2 O.

DEPV:
6. The digested pSV1 vector and pBV with recombination cassette are run on a 1% low melting Seaplaque GTG agarose at 75 V for 8-10 hours. The DNA should be run in a large well created by taping together several teeth of the comb.

DEPV:
7. An 11 kb linearized plasmid band should be visible on the gel for pSV1.RecA. Cut this band and also the recombination cassette insert band from the gel. Purity these DNA fragments using Geneclean Spin columns (Bio 101, Inc.) according to manufacture's direction. Run a small portion of the purified DNA on a gel to estimate the DNA concentration.

DEPV:
8. Ligation reaction: Each ligation reaction is done in 20 ul total volume containing: >50 ng pSV1.vector, 100-200 ng insert, 2 ul 10.times. ligation buffer (Boehringer-Mannheim), 2 ul 10 mM ATP, 1 ul ligase (Boehringer-Mannheim) and ddH2O. Ligation is carried out at 16.degree. C. overnight.

DEPV:
1. Transform the competent BAC cells with the Ts shuttle vector, using 10 ul of the midiprep DNA and 200 ul BAC containing competent cells. Transformation is done as in (IV) of part 11. Plate 1/10 of the transformed cells onto Tet+Chl plates, and grow overnight at 30.degree. C.

DEPV:
4. Make miniprep DNA from a 1.5 ml miniculture using standard alkaline lysis methods. Dissolve the DNA in a 30 .mu.l TE and use 5-10 .mu.l of the DNA for restriction enzyme analysis.

DEPV:
5. Restriction digest with appropriate enzymes and analysis of the co-integrate by Southern blot. Due to the high efficiency of co-integrate formation even with 500 bp homology (>10%), I usually only analyze co-integration on one homology side (either A or B). For example, to analyze co-integrate on A side, use fragment A as a probe and digest the BAC DNA with an enzyme that will detect the co-integrate formation on A side (such as EcoRI). Standard southern blots are done to reveal the co-integrates. As controls, the original BAC and the shuttle vector should be included in this analysis. The reason to use the homology arms as Southern blot probes is that it will hybridize to two bands of appropriate size in the co-integrate BAC. As controls, the original BAC and the shuttle vector should be included in this analysis.

DEPV:
5. A) Two alternative methods can be used to identify the correctly resolved BACs. If both A and B homology are about the same length, one can just pick 10-20 colonies, prepare miniprep DNA by alkaline lysis and do Southern blot to analyze the targeting events. About half of the resolved BACs should contain the correctly targeted marker genes. B) If the two homology arms are not the same length (>500 bp difference), one should use the colony hybridization to select the correctly resolved BACs. Pick

50-100 individual colonies from FA+Chl plates, streak them onto Chl plates and also onto the Tet+Chl plates, as a control for Fusaric acid selection. Each plate can accommodate 50 test colonies and two positive control colonies, which are the co-integrate colonies from the Chl plate. Grow the colonies overnight at 37.degree. C. Abundant colonies should grow on the Chl plate, and none on the Tet+Chl plate, except the positive co-integrate controls. The selection for tet sensitivity at step 4 is very stringent and has essentially no background. Therefore, all the colonies that grow on FA+Chl plates have been found to contain resolved colonies. Colony hybridizations is performed, according to the standard protocols [Sambrook et al., (1989) supra], to select for the colonies that are resolved and resulted in targeted modification. The colony hybridization probe should be part of the recombination cassette excluding the arms, such as lacZ, Neo, GFP or polyA sequences.

DEPV:
6. Midi-prep DNA are prepared for the positive clones by the alkaline lysis method as described above. Restriction digests and Southern blots are performed to confirm targeting event on both homology side (A and B).

DEPV:
1. Digest 50 ug cesium banded BAC maxiprep DNA overnight in 500 ul total volume containing:

DEPV:
4. Now add 5 ul 10.times.DNA dye into the 0.5 ml digested BAC DNA. Take the reservoir out and gently add the DNA(+dye) onto the top of the column with a pasteur pipette. Wait until the DNA+dye just goes into the column, gently add 0.5 ml of injection buffer on top of the column.

DEPV:
7. Purified DNA is stored at 4.degree. C. It is stable for weeks (e.g., no degradation was detected after 3 weeks).

DEPW:
100 .mu.l medi-prep DNA (2-5 ug) or

DEPW:
50 .mu.g DNA

CLPR:
2. The method of claim 1 wherein the recombination deficient host cell cannot independently support homologous recombination because the host cell is RecA.sup.- ; and wherein inducing the host cell to transiently support homologous recombination comprises inducing the transient expression of a RecA-like protein in the host cell.

CLPR:
3. The method of claim 2, wherein inducing the transient expression of the RecA-like protein is performed with a conditional replication shuttle vector that comprises a nucleotide sequence encoding a RecA-like protein; and wherein the expression of the RecA-like protein in the host cell is due to the expression of said nucleotide sequence encoding the RecA-like protein.

CLPR:
5. The method of claim 4 wherein inducing the transient expression of the RecA-like protein comprises:

CLPR:
8. The method of claim 7 wherein neither the IOBCV alone, nor the IOBCV in combination with the host cell, can independently support homologous recombination because both the IOBCV and the host cell are RecA.sup.- ; wherein inducing the host cell to transiently support homologous recombination comprises transiently expressing a nucleotide sequence encoding a RecA-like protein in the host cell; wherein the expression of the nucleotide sequence encoding the RecA-like protein supports homologous recombination in the host cell; and wherein inducing the transient expression of the RecA-like protein is performed with a conditional replication shuttle vector that comprises said nucleotide sequence encoding the RecA-like protein.

CLPR:
10. The method of claim 2 wherein the RecA-like protein is controlled by an inducible promoter; and wherein the transient expression of the RecA-like protein is achieved by the transient induction of the inducible promoter in the host cell.

CLPR:

12. The method of claim 11 wherein the TSSV also contains the recombination cassette, and a first nucleotide sequence that bestows resistance to a host cell containing the TSSV against a first toxic agent, wherein said first nucleotide sequence also can be counter-selected against, and wherein the recombination cassette, said nucleotide sequence encoding the RecA-like protein, and the first nucleotide sequence are linked together on the TSSV such that when the nucleic acid integrates into the particular nucleotide sequence, said nucleotide sequence encodino the RecA-like protein and the first nucleotide sequence remain linked together, but neither said nucleotide sequence encoding the RecA-like protein nor the first nucleotide sequence remain linked to the integrated nucleic acid.

CLPR:
18. The method of claim 14 wherein the RecA-like protein is recA.

CLPR:
20. The method of claim 8 wherein the RecA-like protein is controlled by an inducible promoter; and wherein the transient expression of the RecA-like protein is achieved by the transient induction of the inducible promoter in the host cell.

CLPR:
22. The method of claim 21 wherein the bacterium is an E. coli. bacterium.

CLPR:
23. The method of claim 21 wherein the RecA-like protein is recA.

CLPR:
24. A conditional replication shuttle vector that encodes a RecA-like protein.

CLPR:
28. The TSSV of claim 26 wherein the RecA-like protein is recA.

CLPR:
30. An independent origin based cloning vector that contains a nucleic acid that has been directly modified with specificity by having undergone homologous recombination with a conditional replication shuttle vector in a RecA.sup.- host cell, wherein the conditional replication shuttle vector encodes a RecA-like protein.

CLPR:
39. The method of claim 16 wherein the nucleic acid that binds to the labeled probe is selected from the group consisting of a RNA and a DNA.

CLPV:
(i) transforming the host cell with the TSSV at a permissive temperature, wherein said nucleotide sequence encoding the RecA-like protein is expressed in the host cell and supports the homologous recombination between the nucleic acid and the particular nucleotide sequence; and

CLPV:
(ii) growing the host cell at a non-permissive temperature; wherein the TSSV encoding the RecA-like protein is diluted out.

CLPV:
wherein introducing the recombination cassette into the host cells is performed by transforming the host cells with the TSSV; and wherein inducing the transient expression of the RecA-like protein to support homologous recombination comprises:

CLPW:
(i) incubating the host cells at a permissive temperature in the presence of the first toxic agent and the second toxic agent, wherein transformed host cells containing the TSSV and the BBPAC are selected for; and wherein the RecA-like protein is expressed and a first homologous recombination event occurs between the recombination cassette and the particular nucleotide sequence forming a co-integrate between the TSSV and the BBPAC; wherein a TSSV is either free or part of a co-integrate;

CLPW:
(v) incubating the host cells containing the resolved BBPAC in the presence of the second toxic agent, and a counter-selecting agent; wherein the counter-selecting agent is toxic to host cells containing the first nucleotide sequence and whereby host cells containing said nucleotide sequence encoding the RecA-like protein are removed.